5-2010

The functional role of the Drosophila gypsy insulator in the regulation of gene expression

Hyuck Joon Kang
hkang4@utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Hyuck Joon Kang entitled "The functional role of the Drosophila gypsy insulator in the regulation of gene expression." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Biochemistry and Cellular and Molecular Biology.

Mariano Labrador, Major Professor

We have read this dissertation and recommend its acceptance:

Bruce D. McKee, Jae Park, Ranjan Ganguly, Feng Chen

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

-To the Graduate Council:

I am submitting here with a dissertation written by Hyuck-Joon Kang entitled “The functional role of the *Drosophila gypsy* insulator in the regulation of gene expression” I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Biochemistry, Cellular and Molecular Biology.

Mariano Labrador, Major Professor

We have read this dissertation
and recommend its acceptance:

Bruce McKee

Jae Park

Ranjan Ganguly

Feng Chen

Accepted for the Council:

Carolyn R. Hodges
Vice Provost and Dean of the
Graduate School

(Original signatures are on file with official student records.)

The functional role of the *Drosophila gypsy* insulator in the regulation of gene expression

A Dissertation Presented for

the Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Hyuck-Joon Kang

May 2010

DEDICATION

To my father

Tae-Eon Kang

my mother

Jung-Sook Ko

my brother and sisters

Sung-Jun, Seung-Hee,* and *Do-Hee

ACKNOWLEDGEMENT

I would like to express the deepest appreciation to my advisor, Dr. Mariano Labrador, who from beginning to end always listened carefully to my opinion and ideas, and discussed with me to develop my understanding of and interest in molecular genetics. When I was depressed because of unexpected results from my experiment, he advised me to be disappointed but not depressed and to analyze it to find out what I can learn and obtain from it. Without his guidance and persistent advice, this dissertation would not have been possible.

I would like to thank my committee members, Dr. Bruce McKee, Dr. Jae Park, Dr. Ranjan Ganguly, and Dr. Feng Chen for their critical evaluation as well as encouragement of my research work. From the research work of Dr. Bruce McKee, Dr. Jae Park, and Dr. Ranjan Ganguly, I realized that *Drosophila* is a valuable model organism which is able to provide enormous information for understanding biological development, differentiation, and molecular mechanisms. I would like to show my gratitude to Dr. Ana Kitazono. She gave her kind support in a number of ways, including experimental advice and consultation about my personal problems.

I would like to thank Piedad who regarded our lab members as a family and always tried to cheer up our lab. I'm enormously grateful to Heather who supported me as my friend, lab-mate and English tutor. I would like to thank Shaofei, Cherie, Todd and Srilalitha who all shared my personal joys and sorrows, as well as information about research, with me. I also thank Misty for helping with *in situ* hybridization experiments.

Friendship with our lab members, grown from mutual understanding, consideration, and sometimes severe arguments, is a great asset which I received here.

I owe my deepest gratitude to Dr. Chang-Hwan Park of Hanyang Medical School in Korea, who encouraged and supported me to study in the U.S. I also want to thank my best friend, Hee-Sung Han, and friends in Korea who showed enduring friendship to me.

I would like to especially express thanks to my father and mother. I would not have been here without their support. My father's advice that I should "not forget your original resolution and intention when you started to study in the U.S." aroused me from laziness and effete mind. I always appreciate my mother's love and concern to ask every time I talk with her how my health is and what I have eaten, rather than what I am studying and what I have achieved. In addition, I am grateful to my brother and sisters who have helped with my living in the U.S. in a number of ways.

Lastly, I offer my regards and blessings to all of those who supported me in any respect during the completion of my research project.

ABSTRACT

Chromatin insulators are short DNA sequences that, together with enhancers and silencers, orchestrate gene transcription through DNA-protein interactions in eukaryotic genomes. It has been proposed that insulators operate at the chromatin level by generating functionally independent higher-order chromatin domains. Insulators may maintain the integrity of such domains using two properties: blocking enhancer-promoter interactions and blocking heterochromatin spreading. The *gypsy* insulator of *Drosophila* was identified as a region of the *gypsy* retrovirus responsible for the production of tissue-specific mutations in many genes. The Suppressor of Hairy wing [Su(Hw)] protein contains 12 zinc fingers that specifically bind the *gypsy* insulator. Upon DNA binding, Su(Hw) recruits a second protein, Modifier of Mdg4 67.2 [Mod(mdg4) 67.2], and the interaction of both proteins is required for insulator function *in vivo*. We have found that three different arrays of *gypsy* retrovirus insertions in a *yellow* transgene result in unique *yellow* phenotypes, showing that the enhancer-blocking activity of the *Drosophila gypsy* insulators depends on the relative orientation of the *gypsy* retroviruses on the chromosome. We also observed from transgenic lines with *gypsy* retrovirus or insulator insertions that interaction of insulators may be regulated by active enhancers according to the relative positions of the insulators flanking the enhancers. Moreover, we show that *gypsy* insulators can positively modulate *yellow* activation and result in wild-type levels of expression when placed upstream of

enhancers in *yellow* transgenes in which enhancers are placed out of context by λ-DNA spacers and fail to reproduce the expression levels of *yellow* in wings and body cuticle. Our results provide evidence indicating that this phenomenon is independent of the boundary activity. Genetic analysis using *mod(mdg4)67.2* mutant lines containing *gypsy* retrovirus insertions revealed that the *gypsy* insulator may be placed close to the *yellow* promoter region and be intimately involved in transcriptional activation and repression. Therefore, we suggest that insulators may also function by mediating long range interactions between enhancers and promoters.

TABLE OF CONTENTS

CHAPTER **PAGE**

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Gene regulation by nuclear organization of the genome

The ENCyclopedia Of DNA Elements (ENCODE) and the model organism ENCODE (modENCODE) projects launched in 2003 and 2007, respectively, by the National Human Genome Research Institute (NHGRI) are now underway for the goal to complete the annotation of all genes and other functional elements in the human genome as well as the genomes of *Drosophila melanogaster* and *Caenorhabditis elegans* (Celniker et al., 2009). These large scale projects will provide tremendous and crucial information about gene content and functional elements, but will also raise many questions, such as which regulatory elements control each gene and by which functional mechanism the gene selectivity of the regulatory elements is achieved. The transcription of many genes is accomplished by the harmonious action of regulatory elements such as enhancers, silencers, and insulators (Akbari et al., 2006; Kiefer et al., 2008), and these regulatory elements have frequently can be located up to 100kb from their target genes (Kleinjan and van Heyningen, 2005; Lettice et al., 2003; Nobrega et al., 2003). The complexity of gene regulation in the genome has made it difficult for us to understand the functional relationships between regulatory elements and their target genes.

DNA in eukaryotic genomes is wrapping around a core histone complex consisting of pairs of H2A, H2B, H3A and H4A histones to form an octameric nucleosome. A chromatin fiber is formed by continuous reiteration of this nucleosome unit, and then organized into higher-order structures through higher levels of compaction. As a result, the compact package of the chromatin fiber gives rise to a closed chromatin structure inaccessible to regulatory DNA binding proteins, such as transcriptional regulators (Croston and Kadonaga, 1993). Such higher-order structure is unraveled by ATP-dependent chromatin remodeling complexes or enzymes that can covalently modify the histone tails to allow transcription (Luo and Dean, 1999). As a result, enhancers are able to bring protein complexes containing histone acetyltransferase activity (HAT), methylase activity and kinase activity close to the promoter. These enzymes introduce covalent modifications on the N-terminal tails of histones that subsequently recruit ATP-dependent chromatin remodeling complexes. These complexes act on DNA-histone interactions, exposing the DNA to transcription factors and facilitating the assembly of the transcriptional machinery, which eventually results in transcriptional initiation and elongation (Strahl and Allis, 2000). However, the genetic code and our current knowledge of the epigenetic code are not enough to explain the complexity of patterns of gene expression. It has been suggested that the spatial organization of chromosomes in the nucleus may be important for regulation of gene expression (Dekker, 2008).

A link between the spatial arrangement of chromosomes and gene regulation has been demonstrated through diverse experiments. Studies using *in situ* hybridization

have revealed that chromosomes are organized into specific territories within the nucleus (Borden and Manuelidis, 1988; Cremer et al., 2006; Cremer et al., 1988; Pinkel et al., 1988). Chromatin loops extending out from the chromosome territories were observed when the HoxB gene cluster was activated during differentiation or transcriptional up-regulation of the major histocompatiblity complex (MHC), suggesting that looping out of the chromosome territories is related to decondensation of the chromatin for transcription (Volpi et al., 2000). Moreover, chromosome conformation capture (3C) assays showed that looping out the intervening chromatin allows distant regulatory elements to position closely to their target genes for transcriptional regulation (Kleinjan and van Heyningen, 2005; Spilianakis and Flavell, 2004; Tolhuis et al., 2002). These examples suggest that the formation of chromatin loop domains is intimately involved in gene regulation.

The formation of chromatin loop domains also occurs from the site of attachment of chromatin fibers to the nuclear matrix (Galande et al., 2007). The nuclear periphery is a specific territory where condensed heterochromatin regions are frequently localized (Mirsky and Allfrey, 1960) and silencing of reporter genes may be induced (Andrulis et al., 1998). However, transcriptionally active genes are also positioned at nuclear pore complexes (NPCs) crossing the nuclear envelope (Casolari et al., 2004; Marshall et al., 1997). Structural studies of chromatin organization attached to a matrix or scaffold revealed specific sequences, SARs (scaffold attachment regions) or MARs (matrix attachment regions), that mediate the anchoring of the chromatin fiber to the chromosome scaffold or nuclear matrix, respectively (Cockerill and Garrard, 1986;

Mirkovitch et al., 1984). Interestingly, MARs have been found to frequently co-localize with replication origins (Lagarkova et al., 1998; Razin et al., 1986) and regulatory elements such as enhancers (Bode et al., 2003; Gasser and Laemmli, 1986; Petrov et al., 2006) and insulators (Nabirochkin et al., 1998; Udvardy et al., 1985; Yusufzai and Felsenfeld, 2004). It has been suggested from these observations that spatial access of distantly located regulatory elements through the attachment of MARs to the nuclear matrix may affect gene regulation by increasing the possibility of formation of complexes between the regulatory elements (Razin et al., 2007). Indeed, some MARs have been found to be able to elevate the expression of particular genes in a manner independent of their location (Bode et al., 1995; Kalos and Fournier, 1995; Phi-Van and Stratling, 1996). Nonetheless, the functional significance of the chromatin organization formed by the attachment of MARs/SARs to the nuclear matrix or scaffold still remains unanswered.

Insulators in the eukaryotic genome

Enhancers are *cis*-acting regulatory elements that increase the transcriptional level of linked genes in a manner independent of their orientation and distance. In the eukaryotic genome, enhancers may be located tens of kilobases away from their target genes, but some have been found at distances of up to a megabase from their target gene (Lettice et al., 2003; Nobrega et al., 2003). Although some enhancers, such as the AE1 (autoregulatory element 1) enhancer in *Drosophila,* show preferential interactions for their core promoters (Butler and Kadonaga, 2001), most enhancers appear to be

promiscuous. Therefore, long distances between an enhancer and its target promoter could potentially result in the improper activation of neighboring genes. Nevertheless, spatio-temporal gene regulation occurs precisely during development.

In the eukaryotic genome, transcriptionally active genes are frequently embedded in an environment containing extensive regions of condensed chromatin. For example, polytene chromosomes in *Drosophila* show an alternating pattern of highly condensed bands and relatively decondensed interbands (Urata et al., 1995). Chromatin condensation processes propagate by methylation of neighboring H3 lysine 9 along the chromatin fiber and can affect the expression of neighboring genes (Grewal and Moazed, 2003). For instance, when *white* transgene constructs integrate near heterochromatin in *Drosophila*, position effect variegation (PEV) can be induced, resulting in a variegated eye color. Based on these observations, a fundamental question is how to explain the limited range of the promiscuous enhancer activity and heterochromatin spreading in the genome. Chromatin insulators have been suggested to be one mechanism that genes adopt to restrict these inappropriate positive and negative influences from their surrounding environment through the formation of independent domains (Bushey et al., 2008; Valenzuela and Kamakaka, 2006; Wallace and Felsenfeld, 2007).

Insulators are thought to be transcriptionally neutral *cis*-acting regulatory elements, distinct from enhancers and silencers, which directly affect the transcriptional levels of genes. Insulators can be defined as DNA sequences bound by proteins that exhibit chromatin regulatory properties such as the ability to block promoter-enhancer

communication or prevent heterochromatin from spreading along the chromatin fiber. Insulators have been found from yeast to humans (Geyer and Clark, 2002; West et al., 2002). So far, all characterized insulators identified in yeast have been shown to possess only boundary activity (Donze and Kamakaka, 2001; Fourel et al., 1999). For example, the *ChaI* UAS and $tRNA^{Thr}$ gene restrict the spread of silencing from HML1 and HMR silencers, respectively, and STAR elements located between the X and Y subtelomeric repeats constrain telomeric silencing to limited areas.

Almost all insulators found in vertebrates bind the CTCF protein. These CTCF-binding insulators have been shown to block enhancer activity at several loci and in several species, therefore CTCF has been regarded as a vertebrate enhancer blocking protein. For example, CTCF binds to both the 5' and 3' DNase I hypersensitive sites (HSs) of the chicken β-globin locus. The 5' HS4 and 3'HS protects the β-globin locus from the upstream folate receptor gene enhancer and from the downstream odorant receptor gene enhancer, respectively (Bell et al., 1999). In human and mice, a differentially methylated domain (DMD) located between *H19* and *Igf2* genes acts as an insulator involved in the imprinting of the *H19/Igf2* locus. On the maternally inherited allele, binding of CTCF to the DMD causes insulator activity, preventing the interaction of the *Igf2* promoter with a set of shared enhancers downstream of *H19*. The DMD on the paternally inherited allele is methylated. This methylation blocks CTCF binding and so the insulator activity of DMD is not induced, allowing the *Igf2* promoter to interact with the downstream enhancers (Bell and Felsenfeld, 2000; Hark et al., 2000; Thorvaldsen et al., 1998). Boundary activities have also been found in some CTCF-binding

insulators. For example, the 5' HS4 of chicken β-globin locus acts not only as the folate receptor gene enhancer-blocker but also as the 5' boundary of the open β-globin chromatin structure (Pikaart et al., 1998; Prioleau et al., 1999). However, the boundary activity of the 5'HS4 is independent of CTCF. Recently, the CTCF protein was reported to interact with components of Cohesin, a protein complex that holds together sister chromatids during synthesis through the G2 phase and into mitosis (Parelho et al., 2008; Rubio et al., 2008; Wendt et al., 2008). This interaction indicates that Cohesin is involved in insulator function and strongly supports the notion that Cohesin plays an important role in gene regulation.

In *Drosophila*, the scs and scs' elements are the first DNA sequences described as having the properties of an insulator. The two elements mark the chromatin boundaries of the 87A7 hsp70 locus and contain binding sites for Zw-5 and BEAF proteins, respectively (Gaszner et al., 1999; Zhao et al., 1995). Binding of both proteins is required for the function of the scs and scs' insulators. Fab-7 and Fab-8 insulators were reported to maintain the functional independence of the lab-7 enhancer at the *Abdominal-B (Abd-B)* locus, which is involved in the formation of *D. melanogaster* parasegments during development (Hagstrom et al., 1996; Zhou et al., 1996). Later it was revealed that the dCTCF protein, a counterpart of vertebrate CTCF, binds to most borders between parasegment-specific regulatory domains containing infraabdominal domain (Iab) enhancers in the 50 kb regulatory region (Holohan et al., 2007; Mohan et al., 2007; Smith et al., 2009). Therefore, CTCF is now recognized as a universal insulator protein that is conserved across the phylogenetic scale, including *Drosophila*

and nematodes (Heger et al., 2009; Moon et al., 2005). Also, several insulators containing the binding sites for the GAGA factor protein have been found in *Drosophila*, including the *eve* promoter (Ohtsuki and Levine, 1998) and SF1 boundary (Belozerov et al., 2003). A large part of our current understanding of insulator properties and function actually comes from studies of the *gypsy* insulator (also known as Su(Hw) insulator) of *Drosophila* (Byrd and Corces, 2003; Cai and Shen, 2001; Gerasimova and Corces, 1998; Gerasimova et al., 1995; Ghosh et al., 2001; Muravyova et al., 2001; Roseman et al., 1993; Soshnev et al., 2008).

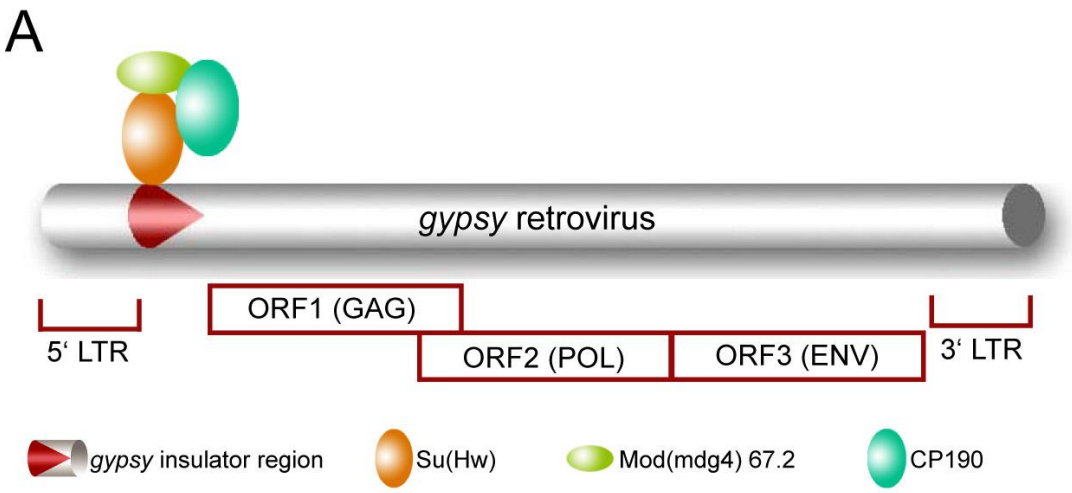
Gypsy insulator and three core insulator proteins

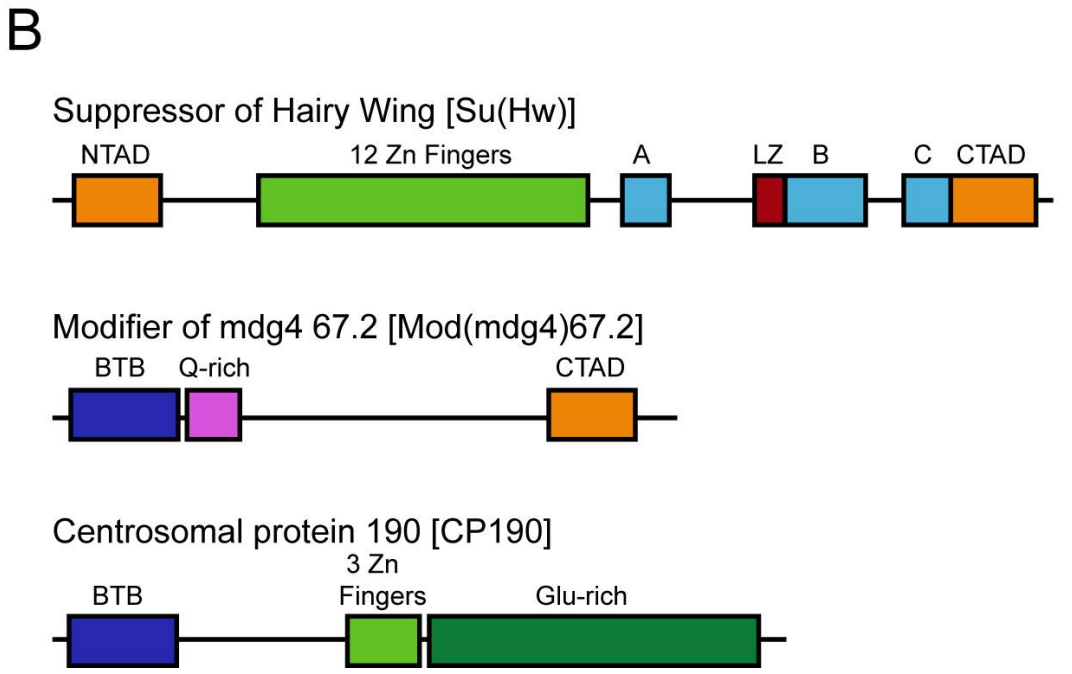
The *Drosophila gypsy* insulator comprises a 350bp sequence downstream of the 5' long terminal repeat (LTR) in the *gypsy* retrovirus (Byrd and Corces, 2003; Gerasimova et al., 2000; Gerasimova and Corces, 1998). The insulator DNA consists of twelve repeats of a motif sequence (5'-YRYTGCATAYBY-3') that binds the Suppressor of Hairy-wing [Su(Hw)] protein and short AT-rich sequences separating the 12 Su(Hw)-binding motifs (Spana et al., 1988). In addition to Su(Hw), at least two other proteins have been found to be required for *gypsy* insulator function: Modifier of mdg4 67.2 [Mod(mdg4) 67.2] and Centrosomal Protein of 190 kD (CP190) (Figure 1.1) (Gerasimova and Corces, 1998; Gerasimova et al., 1995; Pai et al., 2004).

The Su(Hw) insulator protein, encoded by the *su(Hw)* gene located at the 88A12-B2 cytological map position on the third chromosome, consists of 944 amino acids. The Su(Hw) protein contains a stretch of 12 zinc finger motifs at its central portion. Some of

Figure 1.1 *gypsy* retrovirus and three core insulator proteins

(A) The *gypsy* insulator containing 12 Su(Hw) binding sites is illustrated as the red region in the retrovirus. Long terminal repeats and three open reading frames for the gag, pol, and env proteins are depicted below the *gypsy* retrovirus. Ovals of orange, lime and turquoise colors represent Su(Hw), Mod(mdg4)67.2 and CP190, respectively. **(B)** The structures of three core insulator proteins are presented as diagrams. The various domains are depicted as colored boxes; NTAD, amino-terminal acidic domain; LZ, leucine zipper; CTAD, carboxy-terminal acidic domain; BTB, BTB/POZ domain; Q-rich, glutamine rich domain; Glu-rich, glutamate rich domain. A, B, and C in structure of Su(Hw) protein indicate regions highly conserved among *Drosophila* species.


A
gypsy retrovirus
ORF1 (GAG)
5‘ LTR
ORF2 (POL)
ORF3 (ENV)
3‘ LTR
gypsy insulator region
Su(Hw)
Mod(mdg4) 67.2
CP190


B
Suppressor of Hairy Wing [Su(Hw)]
NTAD
12 Zn Fingers
A
LZ
B
C
CTAD
Modifier of mdg4 67.2 [Mod(mdg4)67.2]
BTB
Q-rich
CTAD
Centrosomal protein 190 [CP190]
BTB
3 Zn Fingers
Glu-rich

the 12 zinc finger motifs are involved in recognizing and binding the 5'-YRYTGCATAYBY-3' repeats in the *gypsy* insulator DNA sequence (Harrison et al., 1993; Kim et al., 1996). Acidic domains are found in the N-terminal and C-terminal regions of the protein and a leucine zipper domain is located between the zinc finger domain and the C-terminal acidic domain. Also, sequence comparisons of Su(Hw) proteins from various *Drosophila* species revealed three regions (A, B, and C) which are highly conserved among species but that lack recognizable homology to any functional domain. Regions B and C, which are located between the leucine zipper domain and the C-terminal acidic domain, together with the leucine zipper domain, are involved in the interaction with the Mod(mdg4)67.2 protein (Gdula and Corces, 1997).

Mod(mdg4)67.2 is one of the Mod(mdg4) isoforms encoded by the *mod(mdg4)* gene that is located at the 93D7 cytological map position on the third chromosome. The protein is generated by specific trans-splicing processing to connect exons from two different primary RNA transcripts (Labrador et al., 2001; Mongelard et al., 2002). Mod(mdg4)67.2 contains the highly conserved BTB/POZ domain in the N-terminal region, an adjacent glutamine(Q)-rich domain, and a C-terminal acidic domain through which the protein interacts with the Su(Hw) protein (Gause et al., 2001; Ghosh et al., 2001). Structural data obtained from the PLZF protein demonstrated that BTB domains interact with each other symmetrically to form stable homodimers (Ahmad et al., 1998). Indeed Mod(mdg4)67.2 protein was reported to be able to interact with itself through its BTB domain (Ghosh et al., 2001).

A third *gypsy* insulator protein, CP190, was originally named DMAP190 (*Drosophila* microtubule associated protein of 190 kD) because it was first identified by microtubule affinity chromatography (Jimenez and Goday, 1993). The CP190 protein associates with centrosomes at the onset of mitosis (Jimenez and Goday, 1993; Oegema et al., 1995), but CP190 was found to be dispensable for centrosomal function (Butcher et al., 2004; Pai et al., 2004). CP190 was found to be essential for *gypsy* insulator function through an ethane methyl sulfonate (EMS) mutagenesis screen (Pai et al., 2004). CP190, which is able to interact with both Su(Hw) and Mod(mdg4) 67.2 proteins as well as with itself, consists of a BTB domain at the N-terminal region, a Glu-rich C-terminal region, and three zinc-finger motifs potentially capable of interacting with DNA in its central region. Nevertheless, the consensus DNA binding sequence for CP190 is still unclear.

Endogenous *gypsy* insulators and chromatin loop domains

Su(Hw) and Mod(mdg4)67.2 proteins co-localize at several hundred sites on polytene chromosomes of *Drosophila* larval salivary glands which do not correspond to *gypsy* retrovirus insertion sites (Figure 1.2A and B) (Gerasimova and Corces, 1998). Therefore, these sites are presumed to be endogenous *gypsy* insulators that may be involved in regulating the expression of genes in the *Drosophila* genome. The CP190 protein is located at every site where Su(Hw) and Mod(mdg4)2.2 are co-localized and it also binds at other sites where the Su(Hw) and Mod(mdg4)-67.2 proteins are absent (Pai et al., 2004). In diploid interphasic cells, however, co-localization of the three

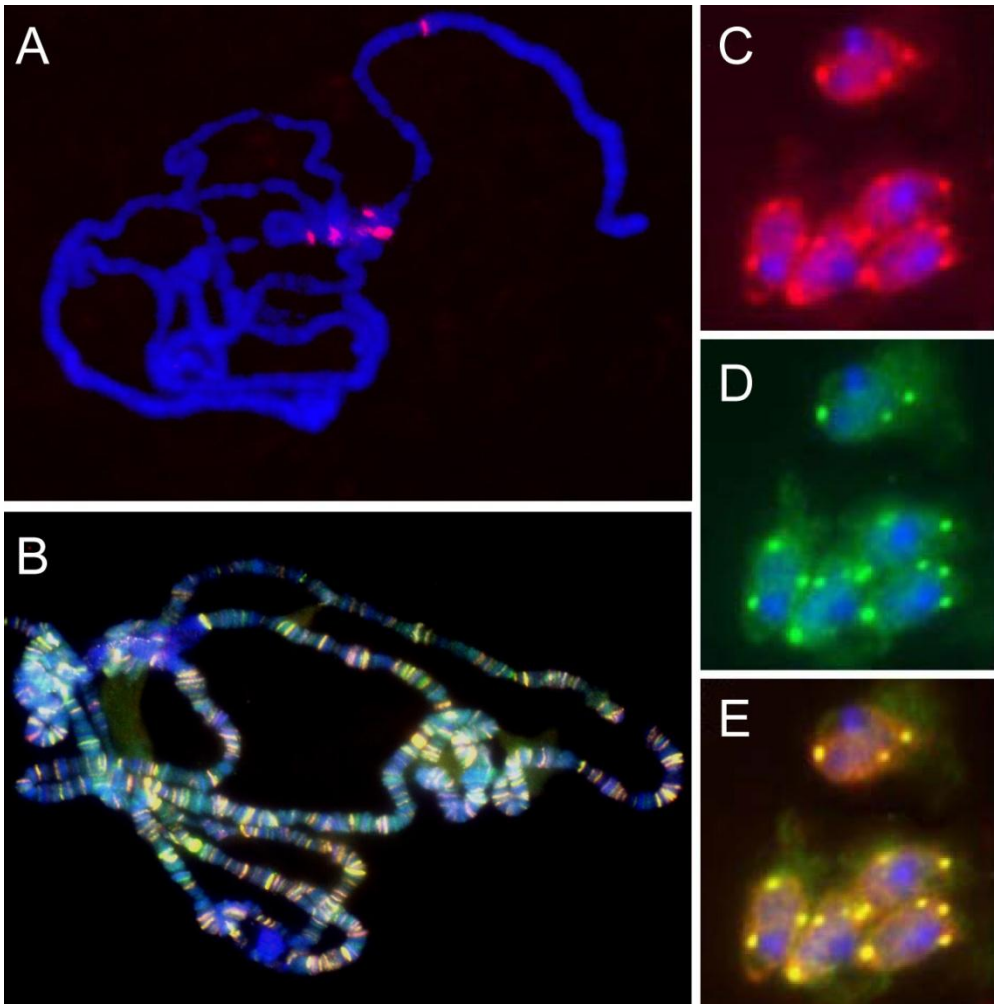


Figure 1.2 Distribution of Su(Hw) and Mod(mdg4) 67.2 on polytene chromosomes and in diploid cells.

(A) *In situ* hybridization of *gypsy* retrovirus (red) on polytene chromosomes. DNA is stained with DAPI [blue]. **(B)** Immunolocalization of Su(Hw) [red] and Mod(mdg4)67.2 [green] on polytene chromsomes. **(C)** Immunolocalization of Su(Hw) [red] and DAPI [blue] in diploid cells. **(D)** Immunolocalization of Mod(mdg4)67.2 [green] and DAPI [blue] in diploid cells. **(E)** Merged image showing co-localization of Su(Hw) (B) and Mod(mdg4)67.2 (C) in diploid cells.

proteins at approximately 500 sites on polytene chromosome is shown as only 25-30 dots per nucleus (Figure 1.2C, D, and E) (Gerasimova et al., 2000; Gerasimova and Corces, 1998; Pai et al., 2004). It has been suggested from these results that coalescence of several endogenous insulators, resulting in a structure called insulator body, triggers the formation of independent loop domains consisting of the intervening DNA sequences (Figure 1.3A and B) (Labrador and Corces, 2002). In a recent study, insulator bodies observed in diploid cells were suggested to be simple aggregates of proteins because removal of the Q-rich domain in Mod(mdg4) prevented the formation of the speckles regarded as insulator bodies but did not affect insulator function, while removal of the Mod(mdg4) C-terminal acidic domain, which interacts with Su(Hw), caused the opposite result (Golovnin et al., 2008). Nevertheless, this model has been supported both directly and indirectly by diverse experimental evidence. For example, *in situ* hybridization as well as nuclear halo techniques revealed that two distal *gypsy* insulators are able to come together in the nuclei of diploid cells in which the insulator proteins anchor to the nuclear matrix (Byrd and Corces, 2003; Gerasimova et al., 2000). Moreover, the insertion of two *gypsy* retroviruses between a promoter and an enhancer led the enhancer to bypass the insulators and activate the promoter, suggesting that the interaction of two insulators may loop out the sequences between them and bring the enhancer and promoter into close proximity (Figure 1.3C) (Cai and Shen, 2001; Muravyova et al., 2001). In addition to the *gypsy* insulator, studies of other insulators have also provided evidence for interaction between insulators. For example, it has been reported that distally located CTCF insulator sites in vertebrates physically interact

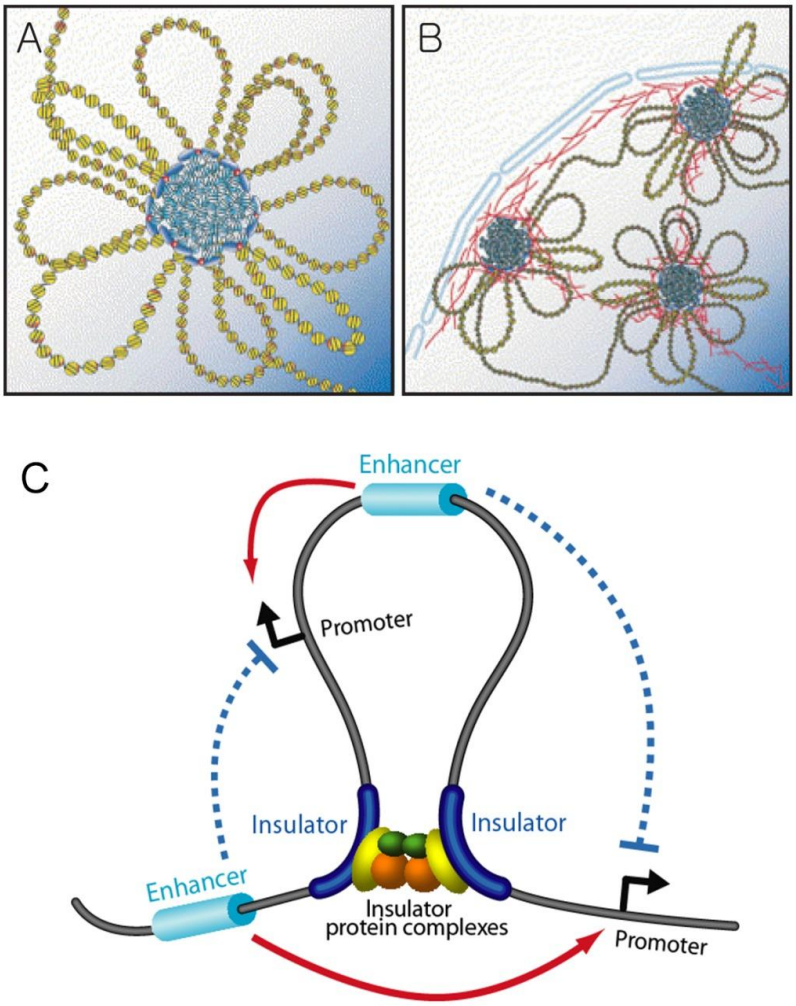


Figure 1.3 Insulator body and a model illustrating the formation of independent chromatin domains mediated by insulators.

(A) Through the formation of insulator bodies, insulators separate the chromatin fiber into loop domains, forming rosette-like structures. **(B)** These might be attached to a fixed perinuclear substrate such as the nuclear lamina (Labrador and Corces, 2002). **(C)** Enhancers outside the domain activate only promoters outside the domain, whereas enhancers within the domain only activate promoters inside the chromatin domain.

in a manner that is dependent on gene activity (Hou et al., 2008; Ling et al., 2006; Splinter et al., 2006). *Drosophila* scs and scs' insulators can also interact with each other via Zw-5 and BEAF proteins bound to each element (Blanton et al., 2003).

The recently identified consensus motif of endogenous Su(Hw) binding sites was revealed to be intimately related to the Su(Hw) binding motif of the *gypsy* insulator (Adryan et al., 2007), suggesting that the endogenous *gypsy* insulators may participate in gene regulation in the *Drosophila* genome through functional insulator properties such as enhancer blocking and boundary function. Indeed, when some endogenous *gypsy* insulator sequences were tested for insulator properties, they acted as insulators to block enhancer activities (Kuhn-Parnell et al., 2008; Parnell et al., 2006; Parnell et al., 2003). The distribution of binding sites of *Drosophila* insulator proteins obtained from genome-wide ChIP-chip analysis supports that heterochromatin boundary and enhancer-blocking activity may be genuine endogenous insulator functions (Bushey et al., 2009; Negre et al., 2010). These analyses also revealed that several insulator proteins co-localize far more often than would be expected by chance. For example, CP190 sites co-localize with 47% of Su(Hw) sites, 62% of dCTCF sites, and 71% of BEAF sites (Bushey et al., 2009) and sites occupied by more than one insulator protein account for 45 % of the identified insulator binding sites (Negre et al., 2010), suggesting that these insulator proteins may have functional implications which are likely to be mediated by CP190. Nevertheless, the distribution of *Drosophila* insulator proteins showed specific differences. BEAF, dCTCF, and CP190 binding sites are enriched close to the transcription start site of genes, whereas many Su(Hw) binding sites are

placed at distances far from genes (Bushey et al., 2009; Negre et al., 2010). The distinct distribution pattern of Su(Hw) implies that *gypsy* insulator proteins may be involved in the regulation of gene expression through an indirect and distinctive mechanism. Tests to determine whether the interaction between the *gypsy* insulator and other *Drosophila* insulators induces enhancer-bypass showed that neutralization of insulator activity did not occur in any pair of insulators (Kuhn et al., 2003; Majumder and Cai, 2003). Therefore, it could be assumed that the *gypsy* insulator protein Su(Hw) participates in gene regulation or chromatin organization by interaction with other insulator proteins as well as by unique functional properties. In addition, detailed analysis of the distribution of Su(Hw) protein on polytene chromosomes using *in situ* hybridization and immunostaining showed that Su(Hw) signals in interbands, which consist of decondensed open chromatin, have properties distinct from Su(Hw) signals in DAPI bands, which correspond to condensed chromatin (Wallace et al., 2009). For example, in the *mod(mdg4)* mutant background, Su(Hw) signals in DAPI bands are significantly decreased or removed, while Su(Hw) signals in interbands are not affected. The difference in the ability of Su(Hw) proteins to associate with their binding sites in the absence of Mod(mdg4) suggests that endogenous insulators may play different functional roles according to their context in the genome.

Other proteins involved in the function of *gypsy* insulator

Three core proteins (Su(Hw), Mod(mdg4) and CP190) which are indispensable for *gypsy* insulator function have been found to interact with several other proteins

(Capelson and Corces, 2005; Kurshakova et al., 2007; Lei and Corces, 2006). The DEAD-box putative RNA helicase Rm62, the *Drosophila* homolog of the human p68 helicase, was revealed to interact with CP190 by using immunoaffinity chromatography with an anti-CP190 antibody (Lei and Corces, 2006). Rm62 is involved in dsRNA-mediated silencing, heterochromatin formation and transposon silencing (Csink et al., 1994; Ishizuka et al., 2002). Genetic analysis revealed that Rm62 negatively regulates insulator activity, whereas other RNAi machinery such as piwi and aubergine positively regulate insulator activity. Mutation of Rm62 or RNAi machinery components does not affect the expression level of the three core insulator proteins or their interaction on polytene chromosomes. However, Rm62 and the RNAi pathway have opposite effects on the assembly or maintenance of insulator bodies in diploid cells. It was suggested from these results that RNAs generated by the RNAi pathway are required for structural formation of the insulator bodies and interaction between Rm62 and these RNAs causes disassembly of insulator bodies (Lei and Corces, 2006).

The *Drosophila* homolog of the human Topoisomerase I-interacting RS protein, dTopors, interacts with Mod(mdg4)67.2 and Su(Hw) (Capelson and Corces, 2005). In addition, dTopors associates with the nuclear lamina. The dTopors protein contains a ring finger domain and belongs to the class of E3 ubiquitin ligases. dTopors was reported to be required for the polyubiqutination of a basic helix-loop-helix transcriptional repressor encoded by the *hairy* gene and the subsequent proteasomal degradation of the Hairy repressor protein (Secombe and Parkhurst, 2004). Although genetic analysis showed that dTopors plays a role as a positive regulator of *gypsy*

insulator activity, dTopors does not ubiquitinate the three core insulator proteins and so it appears that E3 ligase activity is not required for insulator function (Capelson and Corces, 2005). dTopors is assumed to be involved in the establishment of chromatin organization through the attachment of the insulator bodies to the nuclear lamina.

Mod(mdg4)2.2 and CP190 can be post-translationally modified by Small Ubiquitin-like Modifier (SUMO) (Capelson and Corces, 2006). SUMO conjugation is linked to numerous biological processes such as transcription, cell cycle, subnuclear transport, and the maintenance of genome integrity (Dohmen, 2004; Hay, 2005). Disruption of the SUMO conjugation pathway by mutation in the *Ubc9* gene, which encodes the *Drosophila* SUMO E2 conjugating enzyme, or the *Smt3* gene, which encodes SUMO, enhanced the function of a partially inactive insulator. This finding suggests that sumolyation of Mod(mdg4) and CP190 negatively affects the activity of the *gypsy* insulator (Capelson and Corces, 2006). However, it is still unclear how SUMO conjugation is related to the activity of the *gypsy* insulator.

E(y)2, the *Drosophila* homolog of yeast Sus1, interacts with TAF9 which is a component of transcription factor TFIID as well the SAGA/TFTC complex involved in transcriptional co-activation (Georgieva et al., 2001). Also, yeast Sus1 is known as a component of SAGA and of the Sac3-Thp1 complex which is involved in mRNA export (Fischer et al., 2004; Rodriguez-Navarro et al., 2004). Therefore, E(y)2/Sus1 seems to play an important role in regulation of gene expression and in the mRNA-exporting machinery. Immunostaining and genetic analysis using *su(Hw)* and *e(y)2* mutant alleles shows a cytological and genetic interaction between Su(Hw) and E(y)2, and this

interaction could occur through the zinc-finger domain of Su(Hw) Interestingly, the E(y)2 protein is required for the boundary function of the *gypsy* insulator while it does not affect the enhancer-blocking activity of the *gypsy* insulator, suggesting that the boundary activity and enhancer blocking activity of the *gypsy* insulator is achieved through different binding partners of Su(Hw), E(y)2 and Mod(mdg4), respectively (Kurshakova et al., 2007).

The role of orientation in functional interaction of *gypsy* insulators

In some *Drosophila* insulators, the relative orientation of paired insulators plays a key role in insulator function (Kyrchanova et al., 2008; Kyrchanova et al., 2007). An example is the Mcp insulator known as a boundary between the iab-4 and iab-5 regulatory domains in the *Abd-B* complex (Karch et al., 1994; Karch et al., 1985). Transgenic assays using paired Mcp insulators inserted between an enhancer and a promoter in opposite orientation resulted in enhancer-bypass, whereas two insulator in the same orientation blocked the enhancer-promoter communication (Kyrchanova et al., 2007). Similar neutralization of paired insulators in an orientation dependent manner was shown in well-characterized *Drosophila* insulators such as scs, scs', and 1A2 (Kyrchanova et al., 2008). Two Fab-7 insulators in opposite orientation also induced enhancer-bypass (Rodin and Georgiev, 2005) but strengthened the enhancer-blocking activity of insulators in the same orientation (Majumder and Cai, 2003). While these insulators showed neutralization of insulation in a manner dependent on their relative orientation, neutralization of paired *gypsy* insulators is insensitive to their relative orientation. For example, tandem repeats (same orientation) of two *gypsy* insulators

inserted between an enhancer and a promoter also caused activation of the promoter by the enhancer (Kuhn et al., 2003; Majumder and Cai, 2003).

It was recently reported that when two *gypsy* retroviruses were inserted between an enhancer and a promoter, bypass of enhancer activity relied on the relative orientations of the *gypsy* elements (Labrador et al., 2008). Specifically, a 10kb genomic DNA fragment encompassing all the *yellow* gene sequences was used for analysis of interactions between *gypsy* insulators. This fragment includes the intron and all coding and the regulatory sequences necessary to express the *Drosophila* Yellow protein in a tissue specific manner that recapitulates the expression of the endogenous *y* gene. The Yellow protein is involved in the production of pigmentation, and mutations at the *yellow* locus change the coloration of the fly cuticle from brownish-black to yellow (Geyer et al., 1986; Labrador and Corces, 2001). The regulatory sequences that control the expression of the *yellow* (*y*) gene consists on a series of functionally independent tissue-specific enhancers that are required for the expression of the *yellow* gene in wing blades, body cuticle, and bristles (*CasperY2.4;* Figure 1.4A). When a single *gypsy* element integrates between the body enhancer and the promoter, the *gypsy* insulator blocked the wing and body enhancers and the transgenic flies with this type of insertion showed yellow wings and yellow abdomen (***y*^{2}** phenotype) (*CasperY2.4y*^{2}; Figure 1.4B). However, the insertion of a *gypsy* element between the wing and body enhancers allowed the body enhancer to activate the promoter but blocked the wing enhancer, resulting in yellow wings and black abdomen (*CasperY2.4y*^{w}; Figure 1.4C). These transgenic male flies carrying a single *gypsy* insertion were crossed with *flamenco*

Figure 1.4 Examples of single *gypsy* retrovirus insertion in the regulatory region of the *yellow* trangene and its enhancer blocking activity.

(A) Wild-type phenotype of *yellow* gene. The *yellow* gene has at least five enhancers that control its expression in five different tissues. Only three tissue-specific enhancers are described in this figure (wing blade, body, and bristles). The three tissue-specific enhancers activate the *yellow* promoter, causing black color in wings, body and bristles. **(B)** An insertion of the *gypsy* retrovirus between the body enhancer and the promoter blocks both wing and body enhancers **(C)** When positioned between wing enhancer and body enhancer, the *gypsy* retrovirus blocks only the promoter-distal wing enhancer, while the promoter-proximal body enhancer still activates transcription in body cuticle.

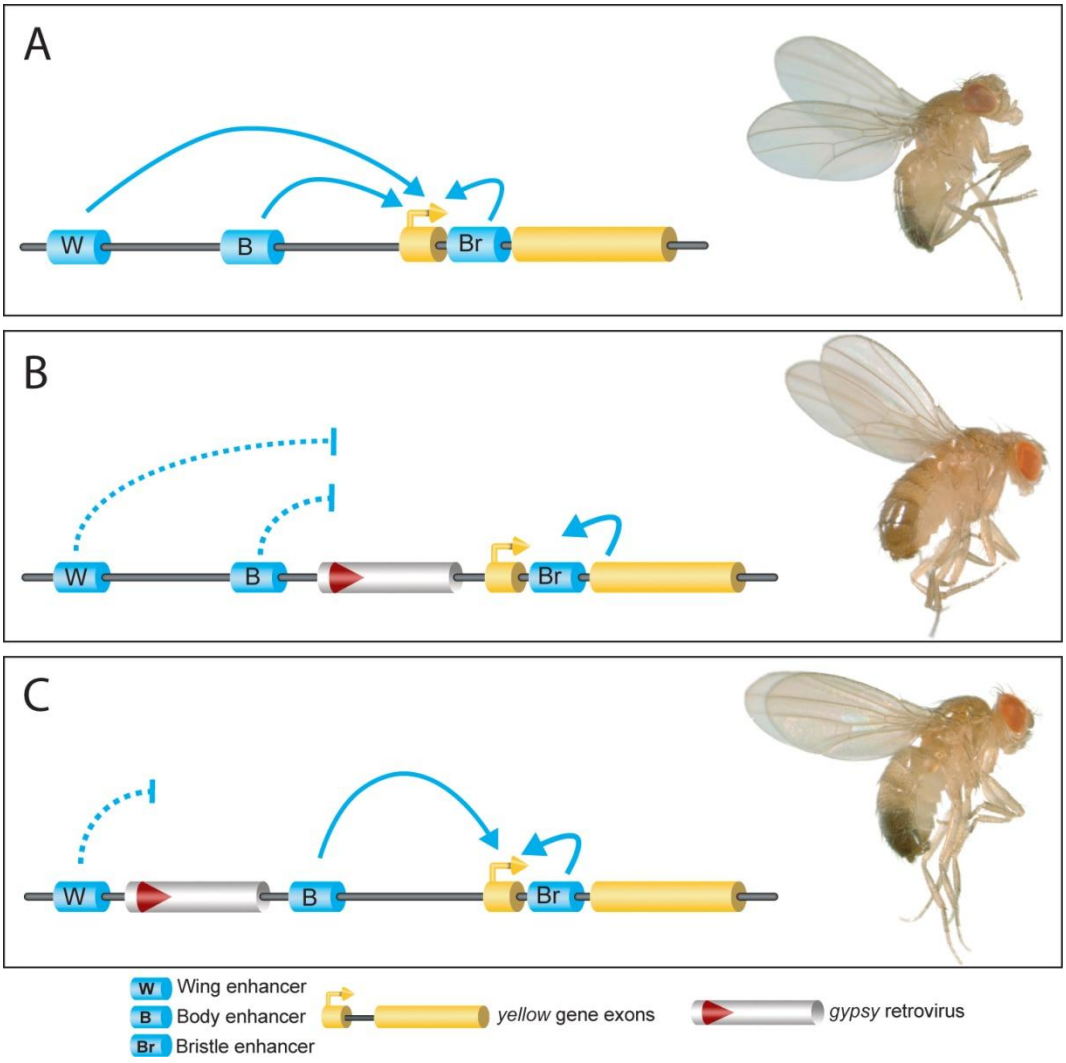
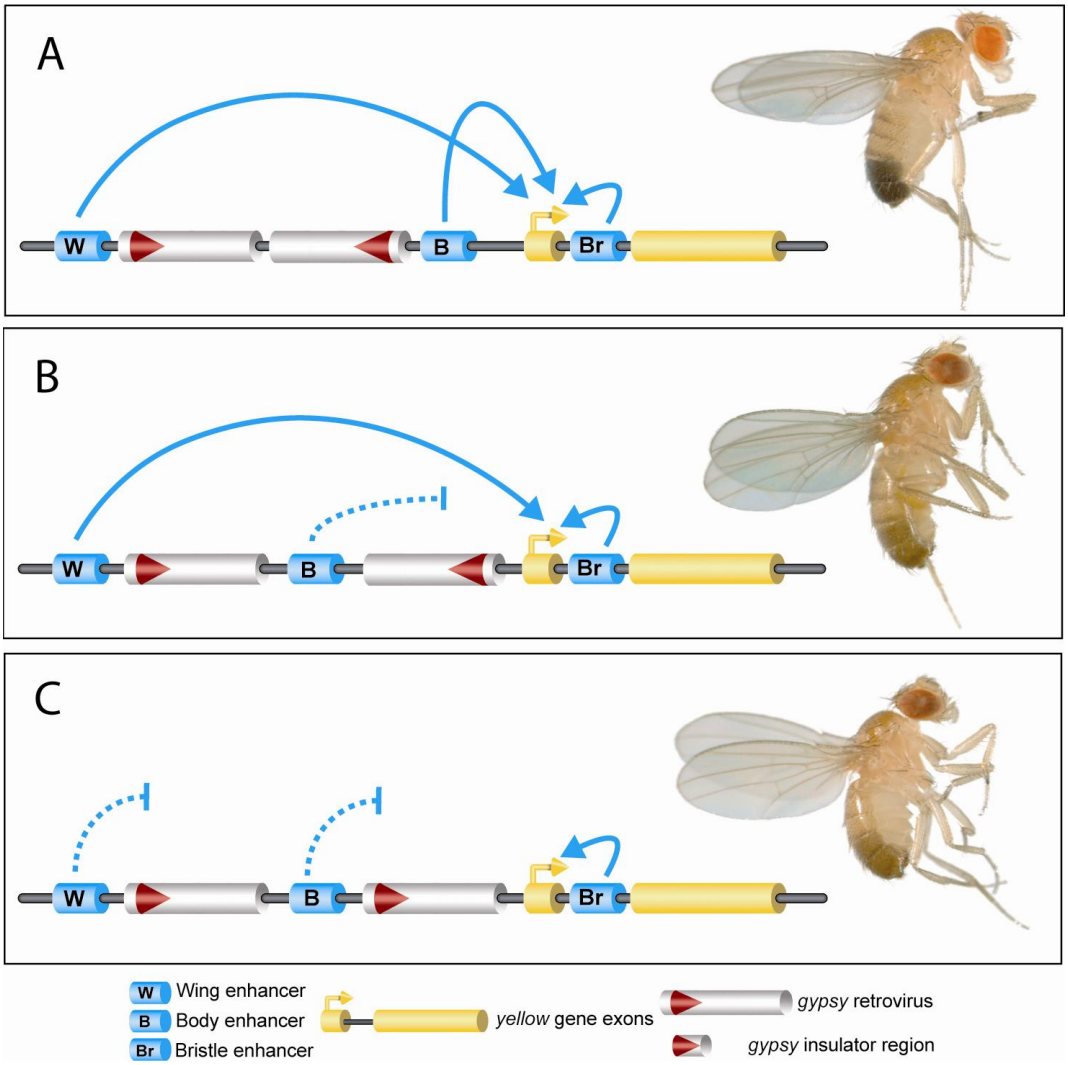

A
W
B
Br
B
W
B
Br
C
W
B
Br
W Wing enhancer
B Body enhancer
Br Bristle enhancer
yellow gene exons
gypsy retrovirus

Figure 1.5 Examples of two *gypsy* retrovirus insertions in the *yellow* transgene and their enhancer blocking activity

When an additional *gypsy* element was inserted into the regulatory region of the *yellow* transgene containing a single *gypsy* retrovirus insertion, changes in the enhancer-blocking activity were induced according to relative position and orientation of the insertions. **(A)** and **(B)** Differences in *gypsy* insertion sites caused changes in enhancer blocking activity and these changes can be explained by the model in Figure 1.3C. **(B)** and **(C)** The relative orientation of two *gypsy* insertions determines changes in enhancer blocking activity that cannot be explained by the model in Figure 1.3C.


A
W
B
Br
B
W
B
Br
C
W
B
Br
W Wing enhancer
B Body enhancer
Br Bristle enhancer
yellow gene exons
gypsy retrovirus
gypsy insulator region

(flam) mutant female flies to induce additional *gypsy* insertions (Dej et al., 1998; Prud'homme et al., 1995). A second *gypsy* element was inserted between the first *gypsy* element and the body enhancer in y^{w} phenotypic flies, which already carried a *gypsy* insertion between wing and body enhancers. This transgenic fly showed a wild-type like phenotype (y^{+}) with black wings and black body, indicating that the wing enhancer is capable of bypassing the two insulators (*CasperY2.4*y^{w+} ; Figure 1.5A). When a second *gypsy* element was inserted between the body enhancer and the promoter in transgenic flies already carrying a *gypsy* insertion between the wing and body enhancers, the phenotype was changed from y^{w} (yellow wings and black abdomen) to black wings and yellow body (*CasperY2.4*y^{wB} ;Figure 1.5B). These results are consistent with the independent loop domain model, which suggests that the physical interaction of two insulators promotes the looping out of the intervening sequence allowing transcriptional activation by the distal enhancer (Figure 1.3C). Similarly, an enhancer flanked by *gypsy* insulators may be looped out, resulting in inhibition of transcription. However, when a second *gypsy* element was inserted between the body enhancer and promoter but the additional *gypsy* element was placed in the same orientation as the first *gypsy* element, the transgenic flies showed an y^{2} phenotype, indicating that the relative orientation of the two *gypsy* insertions may determine changes in enhancer blocking activity (*CasperY2.4*y^{w2}; Figure 1.5C).

Goal of this thesis research

Mutations in genes encoding Su(Hw) and CP190 proteins lead to lethality or to phenotypic malfunctions such as female sterility (Harrison et al., 1993; Pai et al., 2004). The resulting aberrant phenotypes suggest that the formation of functional and independent chromatin domains may be required for cell differentiation and proper somatic and germ line development. Hence, the ultimate role of insulators may be to assist in maintaining the integrity of the eukaryotic genome and orchestrating proper gene expression via the formation of functional and independent chromatin domains. The purpose of this thesis research is to provide genetic and molecular evidence that will help us to understand the mechanism that insulators use to prevent enhancers from interacting with promoters within chromatin domains.

CHAPTER II

MATERIALS AND METHODS

Induction of *gypsy* retrovirus insertions

A third-round *gypsy* insertion was performed by crossing homozygous *y v f mal flam* females to *CasperY2.4* $\boldsymbol{y^{w+}}$ or *CasperY2.4* y^{wB} transgenic males reported in the previous study (Labrador et al., 2008). For the third-round of *gypsy* insertions, heterozygous female offspring from these crosses were individually mated to *y* w^{67c} males. In order to maximize the amount of offspring to be screened, flies were transferred to new vials containing fresh food every 3–4 days. The *yellow* phenotype of the F_2 progeny containing red eye pigmentation was screened under a dissecting microscope, and individuals with new *yellow* mutant phenotypes were identified and used to establish new *Drosophila* lines. Several new mutations were isolated and classified into three new phenotypic classes.

Analysis of *gypsy* insertions by PCR and DNA sequencing

Genomic DNA from 3 to 5 adult flies from each mutant line was extracted using DNAzol (Molecular Research Center, Inc.). Primers *gypsy*-P3 (5'-CTTTGCCGAAAATATGCAATG-3') and *gypsy*-P1 (5'-CAACATGACCGAGGAGCGGTCATAAAC-3'), located outside the LTRs in the 5' and 3'

ends of *gypsy,* respectively, and y0030 (5’-GCCCGATTACCACATTGAG-3’); y1400-(5’-GTTGCACAAAATTACCGGC-3’); y1450 (5’-CTGTGGGTGCAATGATTAG-3’); y1120 (5’-TCATTGCCGCAAGCTCTG-3’); y2900 (5’-CGCCACGGTCCACAGAAGAG-3’), located along the regulatory sequences of the *yellow* gene flanking *gypsy* insertions, were alternatively used to determine the total number of new *gypsy* insertions.PCR amplification of each new *gypsy* insertion sites was also used for further characterization of the insertion site. A combination of P1 with any of the *yellow* primers was employed to detect orientation of the new *gypsy* insertions. The alternative combination of P3 with any of the *yellow* primers detected the same insertions and was used to confirm the results obtained with the first combination of primers. Approximately 50 ng of genomic DNA were used per PCR amplification. Amplified DNA fragments containing *gypsy* insertion sites were directly sequenced by an ABI PRISM 3100 automated DNA sequencer using primers *gypsy*-LTR3 (5’-AAATCGCTATCGCCACAAGGC-3’) or *gypsy*-LTR5(5’-GCAGCGTGAAGCAACACTCCC-3’) located in the 3’ LTR or the 5’ LTR respectively.

In situ **hybridization**

DNA probe was generated from approximately 500 bp DNA fragments obtained by polymerase chain reaction (PCR) using the following sets of primers: WING ENH F (5’- TGTGTAAACGGGAAGTGATC-3’) and WING ENH R(5’-TCTGCCCCAGCACAAAACAA-3’); Yellow F (5’-CACAACTGTCATGTATTAAG-3’) and Yellow R (5’-ATGTCGGCCGCGTTTTATAT-3’). Digoxigenin (DIG) and biotin-labeled

DNA probes were prepared using the DIG DNA labeling Kit and the Biotin High Prime random priming kit, respectively (Roche). The labeled probe DNA was ethanol-precipitated and resuspended in hybridization buffer (4x SSC; 50% formamide, 1x Denhardt, 0.4 mg/ml salmon sperm DNA). Polytene chromosomes obtained from salivary glands were dissected in 0.7% NaCl and transferred for 60 s to 45% acetic acid followed by a 3–4 min fixation in solution 1:2:3 (1 volume lactic acid: 2 volume water: 3 volume acetic acid). Slides were submerged in liquid nitrogen, cover slips were removed, and slides were incubated in 100% cold ethanol, air dried, and stored until hybridization. For hybridization, chromosomes were first pretreated in 2× standard saline citrate (SSC) at 65°C for 30 min followed by dehydration washes with pre-warmed ethanol (twice in 70% for 10 min; twice in 95% for 10 min). Chromosomes were then denatured for 4 min in 0.07 M NaOH and washed for 5 min three times in 2x SSC at room temperature. Chromosomes were then dehydrated at room temperature (twice in 70% EtOH for 5 min; twice in 95% EtOH for 5 min). At this point, probes were denatured for 10 min at 95°C in hybridization solution already mentioned. For hybridization, 30 µl of hybridization solution containing 30 ng of denatured DNA probe placed on the slides were incubated overnight at 37°C in a humidity chamber. Post-hybridization washes consisted of two 2x SSC washes for 10 min at 37°C followed by two washes in 2x SSC for 10 min at room temperature. Slides were then washed twice in phosphate-buffered saline (PBS) for 5 min at RT and once in PBS with 0.1% triton X-100 for 3 min, followed by three times with PBS for 5 min. Blocking was performed by washing twice with 100 mM Tris–HCl and 150 mM NaCl for 3 min, followed by

incubation with blocking reagent (Roche) in 100 mM Tris–HCl and 150mM NaCl for 30min. After blocking, slides were washed twice with PBS for 5 min and once with PBS containing 0.1% NP40 for 5 min. Slides were then incubated with rhodamine-conjugated anti-digoxigenin (1:200 dilution) and Cy5-conjugated Streptavidin (1:150 dilution) at RT for 2 h in PBS containing 0.1% NP40 and 1% nonfat milk. After incubation, slides were washed once in PBS and 0.1% NP40 for 5 min. and twice in PBS for 5 min. Mounting medium containing 4′, 6-diamidino-2-phenylindole (DAPI; Vectashield) was added to the chromosomes, and the hybridization signal was observed using a Leica DM 6000 fluorescence microscope.

Drosophila insulator mutant strains and genetic crosses

Males of genotype y^2wct^6; $mod(mdg4)^{u1}/mod(mdg4)^{u1}$ were mated with *yw;TM6/MKRS* females to remove the y^2 and ct^6 mutant background. F_1 *yw*; $mod(mdg4)^{u1}/TM6$ males were backcrossed to *yw;TM6/MKRS* females. The F2 *yw*; $mod(mdg4)^{u1}/TM6$ strain was selected and maintained. Males of w^{1118}; $su(Hw)^{e04061}/TM6B, Tb^1$ were mated with *yw;TM6/MKRS* females to generate *y* mutant background. F_1 *yw*; $su(Hw)^{e04061}/TM6$ males were crossed to *yw;TM6/MKRS* females. The F2 *yw*; $su(Hw)^{e04061}/TM6$ strain was selected and maintained. To study the influence of $mod(mdg4)^{u1}$ or $su(Hw)^{e04061}$ mutations on the *y* gene expression of the *CasperY2.4* transgenic lines containing one, two, or three *gypsy* retrovirus insertions, the following crosses were performed. The above *CasperY2.4* transgenic lines were crossed to *yw; TM6/MKRS* and then the F_1 *yw;* $P[y^{ovoEN}; w^+]$; *TM6*/+ flies

were backcrossed to *yw; TM6/MKRS*. The F_2 *yw; P[y^{ovoEN}; w^{+}]; TM6/MKRS* flies were crossed to *yw*; *mod(mdg4)^{u1}/TM6* and *yw*; *su(Hw)^{e04061}/TM6* to obtain *yw; P[y^{ovoEN}; w^{+}]; mod(mdg4)^{u1}/TM6* and *yw; P[y^{ovoEN}; w^{+}]; su(Hw)^{e04061}/TM6* flies, respectively. Homozygous *mod(mdg4)*^{u1} or *su(Hw)^{e04061}*mutants were selected from these stocks and the effect of *mod(mdg4)*^{u1} mutants on *yellow* gene expression was observed under a Leica MZ16FA stereomicroscope.

Cloning of pUAS*y* transgenes

The *Drosophila* P element transformation vector pUAST was used to obtain the new pUAS*y* transformation vector. The *HindIII - ScaI* fragment containing the promoter, coding and intron regions of the *yellow* gene was cloned into pBSK. Digestion of this plasmid with *XhoI* and *XbaI* rendered the same fragment, which was subsequently cloned into the multiple cloning site of pUAST. Primers *NotI-y*W (5'-TACTGCGGCCGCTGATTTACCTACCCCT-3') and *AscI-y*W (5'-TTCAGGCGCGCCT-TTACCGGTCTGCAGCCGTTGAACCTAT-3'); *AscI-y*B (5'-ACCTGGCGCGCCAAACCT-AGGCGATCGCATCATTAGTTGCG-3') and *NheI-y*B (5'- AACAGCTAGCTTTCGTA-CGGATTGGATTTCGATTGGGCG-3'); *NheI-y* (5'-TACAGCTAGCAAAGGCCGGCC-GCAATTGAAACGAGCACGAGT-3') and *XhoI-y* (5'-CAATCTCGAGCATGACAGTTGT-GTTCTGAG-3') were used to amplify the wing enhancer, body enhancer and 5' upstream promoter regions, respectively, from the regulatory sequences of the *yellow* gene. Primers were designed with terminal restriction sites in such a way that allowed the cloning of the three PCR amplified fragments consecutively in the following order:

NotI- AscI (containing the wing enhancer) followed by *AscI- NheI* (containing the body enhancer), followed by *NheI - XhoI* (containing the 5' upstream promoter regions of *y*). This strategy generated unique restriction sites *AgeI-AscI-AvrII* in a multiple cloning site between body and wing enhancer as well as *BsiWI-NheI-FseI* in a second multiple cloning site between body enhancer and promoter. Sets of primers *AgeI*-INSL (5'-ACCTACCGGTGTTGTGTATCTGGCCA-CGTAA-3') and *AscI*-INSL (5'-TGAAGGCGCGCCTTGGTTGTTGGTTGGCACAC-3'); *NheI*-INSL (5'-AACAGCTAGCGTTGTGTATCTGGCCACGTAA-3') and *FseI*-INSL (5'-TGTAGGCC-GGCCTTGGTTGTTGGTTGGCACAC-3'); *FseI*-RE-INSL (5'-TACAGGCCGGCCGTTGT-GTATCTGGCCACGTAA-3') and *NheI*-RE-INSL (5'-TGTTGCTAGCTTGGTTGTTGGTT-GGCACAC-3') were designed to clone either one (pUAS*y*-1; Figure 3.11) or two *gypsy* insulators in a head to head orientation as well as in a head to tail orientation (pUAS*y*-2; Figure 3.12). The *gypsy* insulator sequence was amplified from genomic DNA extracted from flies *y v f mal flam*, which contain active *gypsy* elements. The PCR fragment was sequenced, showing that the 12 Su(Hw) binding sites were identical to those found in the *gypsy* insulator sequence used in similar transgenic assays (Roseman et al., 1993).

Cloning of pUAS*y*-λ transgenes

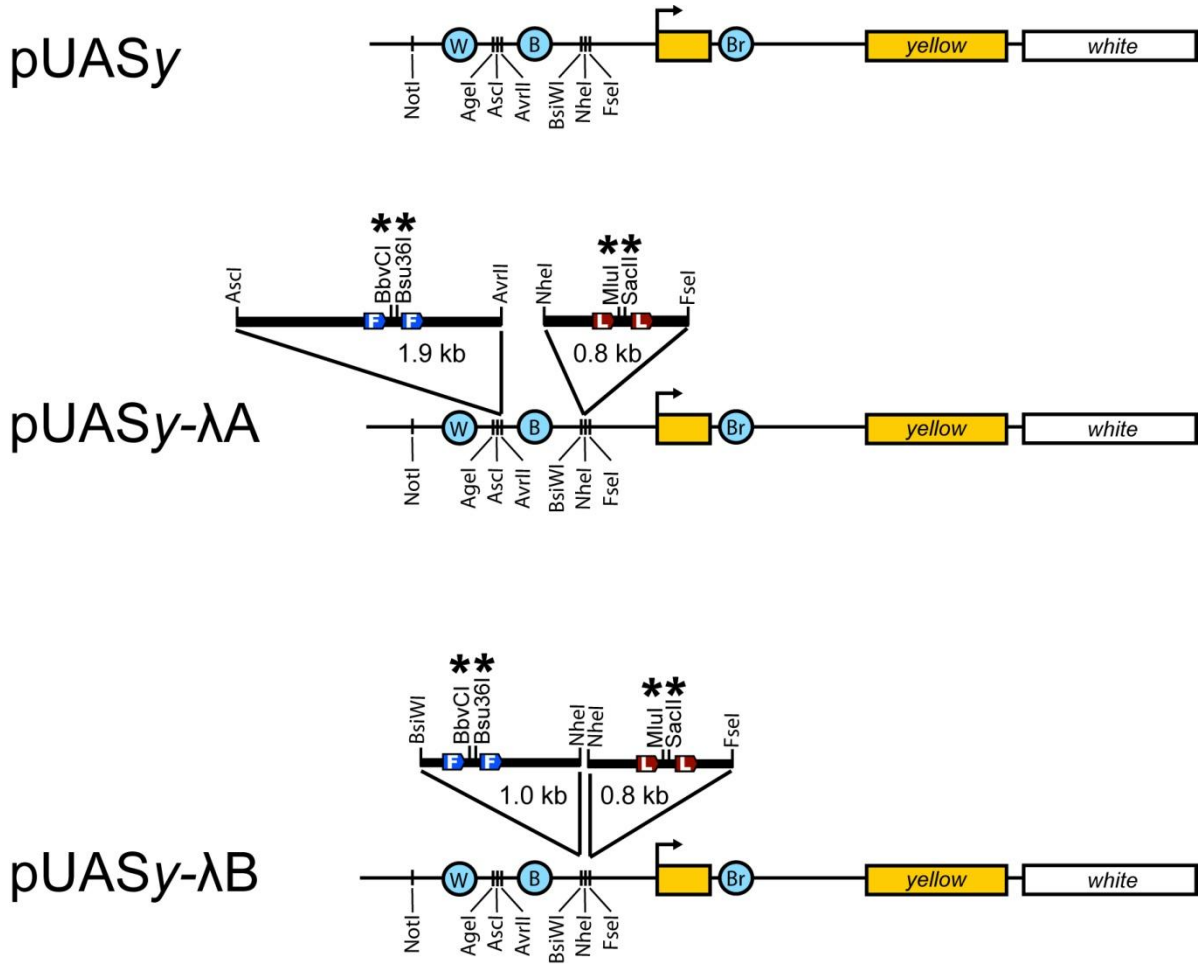
Spacers of 1.9 kb and 0.8 kb were derived from lambda phage DNA. In order to obtain two FRT sites and restriction sites between them for insertion of the insulator fragment, the 1.9 kb lambda spacer was PCR amplified by two primer sets, *AscI*-λ1 (5'-

ACCTGGCGCGCCCGTCTTCCGAAAGAGCTG-3') and *Bsu36I-BbvCI*-FRT-λ1(5'-TAACCTTAGGTGCAGGTCGAGCTGAGGTTAGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCTGGCGGTGACGGTAATTTC-3'); *BbvCI-Bsu36I*-FRT-λ2 (5'-TAACCTCAGCTCGACCTGCACCTAAGGTTAGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCAGATCAACACTGAGCGATGC- 3') and *AvrII*-λ2 (5'-AACACCTAGGGTCTTGCAGACAAACTGC-3'). These two PCR fragments were combined in pCR2.1 vector (Invitrogen), and the combined 1.9 kb *AscI-AvrII* spacer fragment containing *BbvCI* and *Bsu*36*I* restriction sites between two FRT sites was subcloned into *AscI-AvrII* sites between the body and the wing enhancer of the pUAS*y* transformation vector. Two primer sets, *NheI*-λ3 (5'-AAGAGCTAGCTGCGTGTAATTGCGGAGAC-3') and *SacII-MluI*-loxP-λ3 (5'-TATCCGCGGGGCTTAAGCCACGCGTATAACTTCGTATAATGTATGCTATACGAAGTTATGTGATGTCTTCAAGTGGAGC-3'); *MluI-SacII*-loxP-λ4 (5'-TATACGCGTGGCTTAAGCCCCGCGGATAACTTCGTATAATGTATGCTATACGAAGTTATAGAGTCGGCATACAAATATTC-3') and *FseI*-λ4 (5'-TGTAGGCCGGCCGCAAGTATCGTTTCCACC-3') were designed to produce the 0.8 kb *NheI-FseI* spacer fragment for insertion into *NheI-FseI* sites between the body enhancer and promoter in the pUAS*y* vector. The fragment contained *MluI* and *SacII* restriction sites between two loxP sites facing each other for inversion. The resulting pUAS*y*-λA transgene (Figure M1) contained the 1.9kb spacer between wing and body enhancers and the 0.8 kb spacer between the body enhancer and the promoter.

The pUAS*y*-λB transgene containing the 1.8 kb spacer only downstream of the body enhancer (Figure 2.1) was produced by insertion of the 1.0 kb *BsiWI-NheI* fragment and the 0.8 kb *NheI-FseI* fragment into *BsiWI-NheI-FseI* in a second multiple cloning site between the body enhancer and promoter in the pUAS*y* vector. To generate the 1.0 kb *BsiWI-NheI* spacer fragment, the 1.9 kb spacer region was used as a template for PCR amplification with *BsiWI*-λ-1 (5'-AACACGTACGACGGTAACAGCGGCAAC-3') and *NheI*-λ-2 (5'-AAGAGCTAGCGTCTTGCAGACAAACTGC-3') primers. Insulators, amplified with *BbvCI*-INSL (5'-TGTACCTCAGCGTTGTGTATCTGGCCACG-3') and *Bsu36I*-INSL (5'-TGTTCCTTAGGTTGGTTGTTGGTTGGCAC-3') primers and with *MluI*-INSL (5'-TACAACGCGTGTTGTGTATCTGGCCACG-3') and *SacII*-INSL (5'-TGTTCCGCGGTTGGTTGTTGGTTGGCAC-3') primers, were cloned into the *BbvCI-Bsu36I* sites of the 1.9 kb or 1.0 kb spacers and into the *MluI-SacII* sites of the 0.8 kb spacer in the pUAS*y*-λ transgenes, respectively. The insulator upstream of the wing enhancer was produced by PCR with a pair of primers, *NotI*-INSL F (5'-TGTAGCGGCCGCGTTGTGTATCTGGCCACG-3') and *NotI*-INSL R (5'-TGTTGCGGCCGCTTGGTTGTTGGTTGGCAC-3') and cloned into the *NotI* site of the pUAS*y*-λ transgene. The pUAS*y*-λB transgene containing the 1.8 kb spacer only downstream of the body enhancer (Figure 2.1) was produced by insertion of the 1.0 kb *BsiWI-NheI* fragment and the 0.8 kb *NheI-FseI* fragment into *BsiWI-NheI-FseI* in a second multiple cloning site between the body enhancer and promoter in the pUAS*y* vector. To generate the 1.0 kb *BsiWI-NheI* spacer fragment, the 1.9 kb spacer region

Figure 2.1 Schematic representations of pUAS*y*, pUAS*y*-λA, pΔUAS*y*-λA, and pUAS*y*-λB *Drosophila* transformation vectors

Horizontal lines represent the DNA sequence of pUAS*y* transgenic constructs used in this study. The exons of the *yellow* gene and *mini-white* gene are represented as yellow and white boxes, respectively. Blue circles indicate tissue specific enhancers of the *yellow* gene (W-wing blades enhancer, B-body enhancer, and Br-bristles enhancer). Multiple cloning sites are marked as short vertical lines in the *yellow* gene regulatory regions. In the pUAS*y*-λA and pUAS*y*-λB transformation vectors, the thick horizontal lines represent lambda spacer DNA sequences integrated into cloning sites of the pUAS*y* transformation vector. Sizes of each lambda spacer are shown below them (1.9kb, 0.8kb, and 1.0kb). The FRT (F) and loxP (L) target sites for site- specific recombination are drawn as blue and red polygons, respectively, on both sides of the multiple cloning sites in the lambda spacers. Asterisks indicate restriction sites used for insertions of *gypsy* insulator sequences to generate transgenic constructs derived from pUAS*y* and pUAS*y*-λA/B.


pUASy
NotI
AgeI
AscI
AvrII
BsiWI
NheI
FseI
W
B
Br
yellow
white
pUASy-λA
AscI
BbvCI
Bsu36I
AvrII
NheI
MluI
SacII
FseI
F
L
1.9 kb
0.8 kb
pUASy-λB
BsiWI
NheI
1.0 kb
0.8 kb

was used as a template for PCR amplification with *BsiWI*-λ-1 (5'-AACACGTACGACGGTAACAGCGGCAAC-3') and *NheI*-λ-2 (5'-AAGAGCTAGCGTCTTGCAGACAAACTGC-3') primers. Insulators, amplified with *BbvCI*-INSL (5'-TGTACCTCAGCGTTGTGTATCTGGCCACG-3') and *Bsu36I*-INSL (5'-TGTTCCTTAGGTTGGTTGTTGGTTGGCAC-3') primers and with *MluI*-INSL (5'-TACAACGCGTGTTGTGTATCTGGCCACG-3') and *SacII*-INSL (5'-TGTTCCGCGGTTGGTTGTTGGTTGGCAC-3') primers, were cloned into the *BbvCI*-*Bsu36I* sites of the 1.9 kb or 1.0 kb spacers and into the *MluI*-*SacII* sites of the 0.8 kb spacer in the pUAS*y*-λ transgenes, respectively. The insulator upstream of the wing enhancer was produced by PCR with a pair of primers, *NotI*-INSL F (5'-TGTAGCGGCCGCGTTGTGTATCTGGCCACG-3') and *NotI*-INSL R (5'-TGTTGCGGCCGCTTGGTTGTTGGTTGGCAC-3') and cloned into the *NotI* site of the pUAS*y*-λ transgene. For the insulator between the tandem repeats of loxP sites upstream of the wing enhancer (pUAS*y*-λA9), a region encompassing the insulator and loxP sites in pUAS*y*-λA1 transgenes was PCR amplified with two primer sets, *MluI*-INSL (5'-TACAACGCGTGTTGTGTATCTGGCCACG-3') and *NotI*-loxP R (5'-TCAGCGGCCGCATAACTTCGTATAGCATAC -3') primers; *NotI*-loxP F (5'-TCTGCGGCCGCATAACTTCGTATAATGTATGCTATACGAAGTTATACGCGTGTTGTGTATCTGGCCACGT-3') and *SacII*-INSL (5'-TGTTCCGCGGTTGGTTGTTGGTTGGCAC-3'). These two PCR fragments were combined in the pCR2.1 vector (pCR2.1-*NotI*-loxP-INSL-loxP-*NotI*), and the resulting

insulator fragment containing tandem repeats of loxP sites on both sides was inserted into *NotI* sites upstream of the wing enhancer of the pUAS*y*-λA transgene.

Generation and analysis of transgenic lines

The pUAS*y* and pUAS*y*-λ-derived transformation vectors were microinjected into y^{w67c} embryos and all transgenic flies were maintained at 25 °C and 70% humidity on standard corn meal and agar medium. The lines with excisions (pUAS*y*-λA9 Cre1, 2, and 3) or inversions (pUAS*y*-λA4R) were obtained by crossing with lines expressing Cre recombinase (y^1 w^{67c23}; *noc[Sco]*/*Cy*O, *P{w[+mC]=Crew}DH1*). Individual offspring from this cross were transferred to new vials containing fresh medium and crossed with y^{w67c} flies. Three different excisions between the loxP sites upstream of the wing enhancer and loxP sites in the 0.8 kb spacer downstream of the body enhancer were generated by *Cre*-recombinase. The levels of excision in the F2 progeny containing the pUAS*y*-λ transgene were analyzed by PCR with the combination of a reverse primer (Pry1) located at the 3' end of the P element (5'-CCTTAGCATGTCCGTGGGGTTTGAAT-3') and one of two different primers at the following sites: between the insulator upstream of wing enhancer and body enhancer (5'-CTCCCGAAATCGCAACCATAAG-3') and between the 0.8 kb spacer and promoter (5'-GCATAGAATGCATTGCTCTCC-3'). The inversions by Cre recombinase were analyzed with the *gypsy* P3 (5'-CTTTGCCGAAAATATGCAATG-3') primer and primers flanking the insulator insertion site (5'-CAGGCCAACATTGAGTT-3' and 5'-GCAGGTGTTTCCTTGTATAC-3'). The level of expression of the *yellow* gene was estimated by observing the degree of

pigmentation in the body cuticle and wing blades using a Leica MZ16FA stereomicroscope. The levels of pigmentation of 3- to 4-day-old males were scored on a five point scale independently by at least two investigators. A score between 1 and 5 indicates the grade of *yellow* pigment in y^1 mutant and wild-type flies, respectively. Pictures of transgenic flies were obtained using a Leica DFC420 digital camera. This phenotypic analysis of *yellow* expression is standard among laboratories using the *yellow* gene of *Drosophila* and has been reliably used for many years as a tool to estimate transcription levels at the *yellow* gene in the presence of insulators (Kuhn-Parnell et al., 2008; Kuhn et al., 2003; Kyrchanova et al., 2008; Maksimenko et al., 2008; Parnell and Geyer, 2000; Savitskaya et al., 2006).

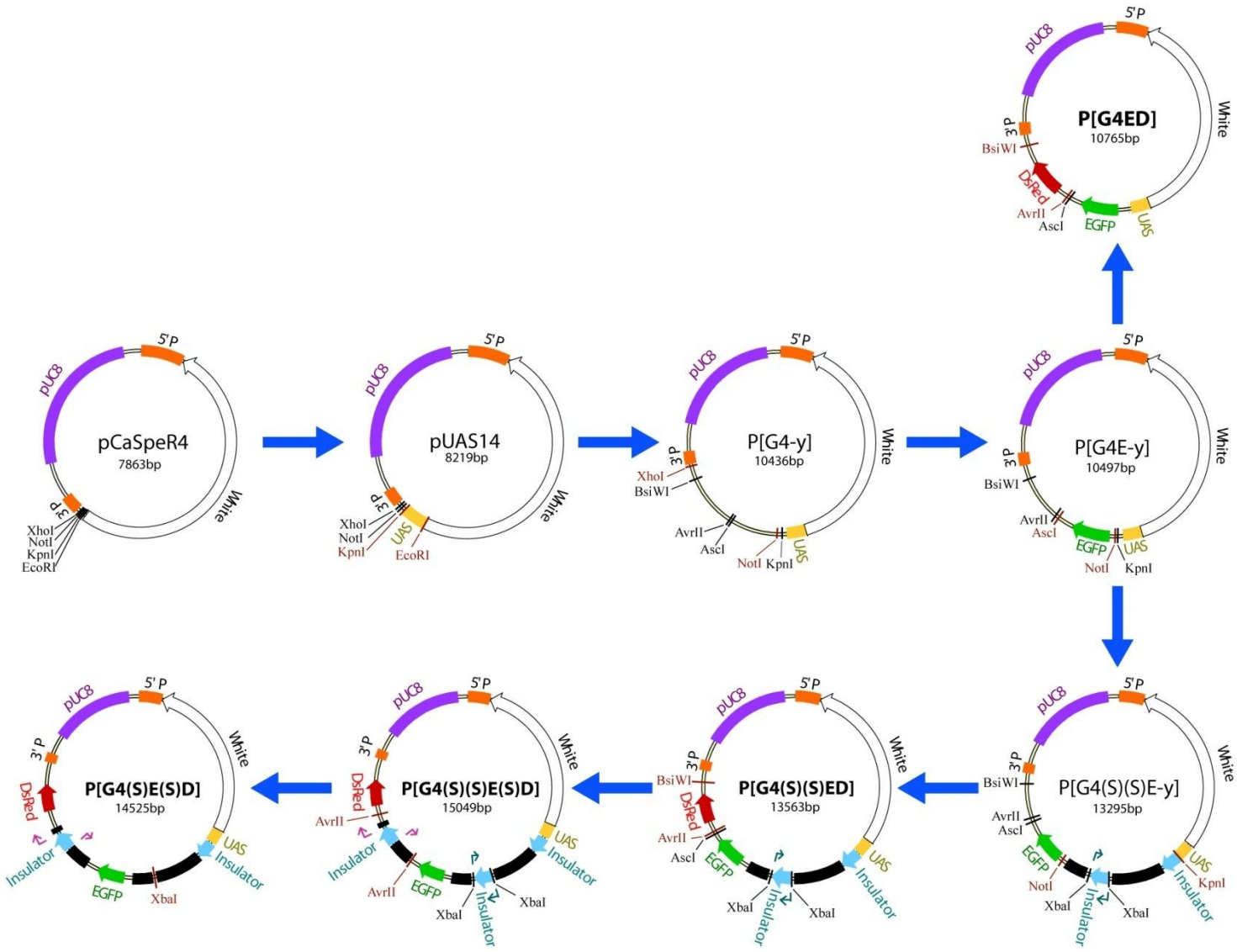
Cloning of GAL4-UAS driven reporter transgenes and Gal4 lines

The *Drosophila* P element transformation vector pCaSpeR4 was used to obtain the pUAS14-EGFP-DsRed transformation vector. The *EcoRI-KpnI* 14X UAS fragment was obtained by PCR amplifying genomic DNA of a EP line with a primer set, *EcoRI* 14XUAS (5'-CATGAATTCCAAGCTTAGGCCTCCAAGGCG-3') and 14XUAS *KpnI* (5'-TCTCTCGGTACCCTCGAC-3'). The *EcoRI-KpnI* 14X UAS fragment was cloned into the pCaSpeR4 transformation vector (pUAS14). The pUAS*y* plasmid described above was digested with *NotI* and *XhoI*, and then the 2.3 kb *NotI-XhoI* fragment was cloned into the pUAS14. The insertion of the 2.3 kb *NotI-XhoI* fragment created the new *BsiWI-NheI-FseI* and *AgeI-AscI-AvrII* multiple cloning sites in pUAS14 plasmid. The following PCR fragments such as reporter genes, *gypsy* insulators, and spacer lambda DNA

sequences were cloned into the new multiple cloning sites of pUAS14 plasmid. The 2.8 kb PCR fragment containing the lambda spacer and two insulator sequences was prepared by PCR amplifying from the pUAS*y*-λA5 plasmid with the two primer sets, *KpnI* I-FI (5'-TTAGGGTACCAATTATAGGTTCAACGGCTGCAG-3') and MLJ-P2 (5'-TAACCTTAGGTGCAGGTCGAGCTGAGGTTAAGCCGGCGCACCGAAG-3'); MLJ-P7 (5'-TGTACCTCAGCGTTGTGTATCTGGCCACG-3') and *NotI* I-FI 2 (5'-CTCGCGGCCGCGTCTTGCAGACAAACTGCGCAA-3'). These two PCR fragments were subcloned in the pCR2.1 vector and then combined to generate the 2.8 kb fragment containing the lambda spacer and two insulator sequences. The combined *KpnI*-*NotI* 2.8 kb fragment was cloned into the pUAS14 vector digested with *KpnI* and *NotI.* The 1.5 kb PCR fragment containing the other lambda spacer and single insulator sequences was amplified with the primer set *AvrII* LI F 2 (5'-AAACCTAGGCATGACAGTTGTGTTCTGAGAACA-3') and *AvrII* LI R 2 (5'-TAACCTAGGTGCGTGTAATTGCGGAGACTTTG-3'). The pUAS*y*-λA1 plasmid described above was used as a template for this PCR amplification. The 1.5 kb *AvrII*-*AvrII* PCR fragment was cloned into the *AvrII* restriction site of the pUAS14 vector. To generate the 1.2 kb EGFP reporter gene fragment containing *NotI* and *AscI* sites on the ends, the pGreen H-Pelican vector was used as template for PCR amplification with *NotI*-EGFP (5'-AATTGCGGCCGCGAGCGCCGGAGTATAAATAGA-3') and EGFP-*AscI* (5'-CGTGGCGCGCCATTAACGCTTACAATTTACGCCTT-3')**.** The PCR fragments of EGFP reporter genes were subcloned into the pCR2.1 vector (pCR2.1-*NotI*-EGFP-*AscI* reverse). The 1.2 kb DsRed reporter gene fragment containing *AvrII* and *BsiWI* sites on

Figure 2.2 Cloning scheme for constructs consisting of 14XUAS, EGFP, DsRed, insulator and lambda spacer

Names and sizes of constructs are shown in plasmid map circles and transgene constructs used for experiments are shown in bold text. In pCaSpeR4, the *white* gene and pUC8 region are shown as white arrow bars and violet bars, respectively. 3'P and 5'P (orange bars) indicate P-elements. EGFP, DsRed and insulators are shown as green, red, and azure arrow bars, respectively. Black bars represent lambda DNA spacers and yellow bars indicate 14X GAL4 binding sites (14X UAS). Representative restriction sites used for cloning are marked. In each plasmid map, restriction sites used for insertion of new DNA fragments are shown in red. Cloning procedure is indicated by blue arrows between plasmid maps.


pCaSpeR4
7863bp
pUAS14
8219bp
P[G4-y]
10436bp
P[G4E-y]
10497bp
P[G4ED]
10765bp
P[G4(S)(S)E-y]
13295bp
P[G4(S)(S)ED]
13563bp
P[G4(S)(S)E(S)D]
15049bp
P[G4(S)E(S)D]
14525bp
pUC8
White
5'P
3'P
UAS
EGFP
DsRed
Insulator
XhoI
NotI
KpnI
EcoRI
BsiWI
AvrII
AscI
XbaI

the ends was amplified by *AvrII* DsRed (5'-TATCCTAGGAGCGCCGGAGTATAAATAGAG-3') and DsRed *BsiWI* (5'-GTGCGTACGCATTAACGCTTACAATTTACGCC-3') primer set with pRed H-Pelican as template. The EGFP and DsRed reporter gene fragments were cloned into *NotI-AscI* and *AvrII-BsiWI* restriction sites of pUAS14, respectively. Each transposon was named to indicate the identities of 14X UAS, the reporter genes and the *gypsy* insulator in the order of arrangement. For example, the P[G4ED] transgene contains 14X UAS[G4], EGFP[E] and DsRed[D] reporter genes. In P[G4(S)(S)ED], two *gypsy* insulators [(S)] were placed between 14XUAS and two reporter genes. In P[G4(S)E(S)D], a single insulator was placed between 14XUAS and EGFP and between EGFP and DsRed. In P[G4(S)(S)E(S)D], two insulators were placed between 14X UAS and EGFP and one insulator was placed between EGFP and DsRed. P[G4(S)E(S)D] was produced from removal of insulator sequences by *XbaI* digestion and self-ligation of P[G4(S)(S)E(S)D] (Figure 2.2). The *GMR*-GAL4 and *Crz*-GAL4 lines were obtained from Dr. Jae Park. Females of the *GMR*-Gal4 lines, which drive eye-specific expression, were crossed with males carrying the test constructs. The eye-specific expression of EGFP and DsRed reporter genes in the pupal stage were observed under a Leica MZ16FA stereomicroscope. Pictures were obtained under the same exposure using a Leica DFC420 digital camera. The Crz-GAL4 lines were crossed with reporter lines and the progeny larvae were dissected to obtain central nervous system (CNS). The CNS tissues were fixed in PBS containing 4% paraformaldehyde for 30 min at RT and then washed three times in PBS for 5 min. The CNS tissues were cleared in PBS containing

30% and 60% glycerol for 10 min. Mounting medium (Vectashield) was added to the CNS sample slides and the EGFP and DsRed signals were observed using a Leica DM 6000 fluorescence microscope. Four or five independent lines containing each reporter transgene were crossed with *GMR*- and *Crz*-GAL4 lines. At least 5 specimens from each independent line were analyzed.

Cloning of p*CI-OMB* transgenes

The ~2.2 kb genomic region containing the *cubitus interruptus* (*ci*) distal regulatory element was PCR amplified by two primer sets, *ci* EN F new (5'-CAACTAGTGTAACATGTCTAAGACCACAGGA) and *ci* EN MR (5'-CTCTTCTTTGGTTCAGTAC-3'); *ci* EN MF (5'-CACAAGTGTGCGAGCATATC-3') and *ci* EN R new2 (5'-CGTTGACTAGTCTTGGCAATG-3') and these two PCR fragments were combined in the pCR2.1 vector (pCR2.1-*ci* enhancer). The ~2.2 kb genomic region is numbered 9240-11411 in the GenBank sequence (accession number U66884.1) and the *cubitus interruptus* (*ci*) distal regulatory element region corresponds to a~1.3 kb fragment numbered 10018-11411. The 2.2 kb fragment containing the *optomotor-blind* [(Secombe and Parkhurst) wing enhancer was PCR amplified by the primer set *OMB* W LF (5'-AAAGGTACCTTTACGCGTAGCAGCGTACTTCGTACTTCGT-3') and *OMB* W RR (5'-CAATCTAGAAATCCGCGGTGGCATAACCAAAACAAGAAATATGCCGA-3'). The 2.2 kb *MluI-SacII* PCR fragment containing the *optomotor-blind* [(Secombe and Parkhurst) wing enhancer was subcloned into pCR2.1-*NotI*-loxP-INSL-loxP-*NotI* (dscribed above), replacing the insulator sequence with the *OMB* wing enhancer

(pCR2.1-*NotI*-loxP-*OMB*-loxP-*NotI*) and the resulting *OMB* wing enhancer fragment contained tandem repeats of loxP sites on both sides. The *optomoto- blind* (Secombe and Parkhurst) wing enhancer upstream of the *bi* (*bifid*) gene, corresponded to a ~1.5 kb fragment, numbered 1-1474 in the GenBank sequence (Accession Number AF291717). The *gypsy* insulator, containing 12 Su(Hw) binding sites and specific restriction sites on the ends, was amplified with three different primer sets creating *NheI-NheI, FseI-KpnI,* and *NotI-NotI* restriction sites on the ends of PCR fragments of insulators. These insulator fragments were subcloned into the pCR2.1 vector. Three different lambda spacer fragments (1.5kb *KpnI-NotI* spacer, 1.7kb *NotI-NheI* spacer, and 400bp *AvrII* spacer) were prepared by PCR amplification with the following primer sets, Lambda A F (5'-TTAGGTACCACCAGGAATGTAGTGGCGGA-3') and Lambda A R (5'-AATGCGGCCGCACCTGAGCTTAGAACCTTTAC-3'); *NotI*-lambdaB (5'-ATTGCGGCCGCTGCTGCATGCGAAAGTCCTACGGTCAAGAG-3') and lambdaB-*Nhe*I (5'-TTAGCTAGCAGATGCTGATATATTTTAGAGGTGA-3'); Lambda C F (5'-TTTCCTAGGCAGGACAACTTCTGGTCCGGT-3') and Lambda C R (5'-ATACCTAGGCTGTTCCATCGTGGTGATCCCGTT-3'). The DsRed reporter gene fragment was prepared by PCR amplification with two primer sets, *SphI*-F-*AvrII*-R (5'-TTCGCATGCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCTATCCTAGGAGC GCCGGAGT-3') and DsRed *BsiWI* (5'-GTGCGTACGCATTAACGCTTACAATTTACGCC-3'); *Avr*II-DsRed (5'-TATCCTAGGAGCGCCGGAGTATAAATAGAG-3') and *SphI*-F-*BsiWI*-R (5'-ATTGCATGCTGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGTGCGTACGCATT

AACGCTTAC-3'). The *AvrII*-DsRed-*BsiWI* fragment, described above, was used as a template for the PCR amplification. The two PCR products were combined in the pCR 2.1vector (pCR2.1-*SphI*-FRT-DsRed-FRT-*SphI*) and this led to the formation of tandem repeats of loxP sites on both sides of the DsRed reporter gene. The 400bp *AvrII* lambda spacer was then subcloned into the *AvrII* site upstream of promoter region of DsRed (pCR2.1-*SphI*-FRT-λC-DsRed-FRT-*SphI*).

The p*CI-OMB* series of constructs (p*CI-OMB* RG, 1, and 2) were made using the pCaSpeR4 vector. The pCR2.1-*NotI*-EGFP-*AvrII* reverse construct described above was digested with *EcoRI* and the digested fragment containing the EGFP reporter gene was inserted into the pCaSpeR4 vector digested with *EcoRI* (pCaSpeR4-EGFP). The 1.7kb *NotI-NheI* lambda spacer was subcloned into pGreen H-Pelican and then the construct was cleaved by *NotI* and *AgeI*. The *NotI-AgeI* fragment containing the 1.7kb lambda spacer was inserted into the pCaSpeR4-EGFP (pCaSpeR-λB-EGFP). Insertion of the spacer fragment created *SphI* and *AvrII* cloning sites and also replaced the *NotI* site adjacent to the promoter region of the EGFP reporter gene with the *NheI* site (pCaSpeR-λB-SG). The *NheI-NheI* insulator fragment was cloned into the *AvrII* site in the lambda spacer. The pCR2.1 vector containing the *FseI-KpnI* insulator fragment was digested with *SpeI* and *NotI*. The *SpeI-NotI* insulator fragment was subcloned into pCaSpeR-λB-SG, creating a new *KpnI* site (pCaSpeR-S-λB-SG). The 2.2kb fragment containing the *cubitus interruptus* enhancer (*ci*) and the 1.5kb lambda spacer were inserted into the *SpeI* site and *KpnI/NotI* sites, respectively (pCaSpeR-*ci*-S- λAB-SG). The *NotI-NotI* fragment containing loxP-*OMB*-loxP was subcloned into the *NotI* site

(pCaSpeR-*ci*-S-L*OMB*L-SG) and then the fragment containing FRT-λC-DsRed-FRT was inserted into pCaSpeR-*ci*-S-L*OMB*L-SG cleaved with *AvrII*. This construct was named p*CI-OMB* RG. The p*CI-OMB* 1 and 2 constructs were derived from the intermediate products for p*CI-OMB* RG.

In p*CI-OMB*1, the fragments containing the *cubitus interruptus* (*ci*) enhancer and FRT-λC-DsRed-FRT were inserted into the *SpeI* and *SphI* sites of pCaSpeR-λB-EGFP, respectively (pCaSpeR-EGFP-*ci*-DsRed). The fragment of pCaSpeR-*ci*-S-L*OMB*L-SG cleaved with *SpeI* and *AvrII* was inserted into the pCaSpeR-EGFP-*ci*-DsRed cleaved with *AvrII*. In *pCI-OMB* 2, The *NotI-NotI* insulator fragment was subcloned into the pCR 2.1 vector (pCR2.1-*NotI*-INSL-*NotI*). One of two *EcoRI* sites in this construct was removed and then the *cubitus interruptus* (*ci*) enhancer fragment cleaved with *SpeI* and *MfeI* was inserted into the *EcoRI-SpeI* sites of the construct (pCR-*ci*-INSL). The *SpeI-XbaI* fragment from pCR-*ci*-INSL was subcloned into pCaSpeR4 cleaved with *XbaI* (pCaSpeR-*ci*-INSL). The fragment containing FRT-λC-DsRed-FRT was cloned into the *SphI* site of pCaSpeR-*ci*-INSL (pCaSpeR-*ci*-INSL-DsRed). The *EcoRI-SacII* fragment and *SpeI-SacII* fragment from the pCaSpeR-*ci*-S-L*OMB*L-SG was subcloned into the pCaSpeR-*ci*-INSL-DsRed, in order of precedence. Sequences of primers, DNA fragments and cloning procedure for the p*CI-OMB* series of constructs are described in detail in Table 2.1, 2.2, and 2. 3. All the p*CI-OMB* constructs were microinjected into yw^{67c} embryos. All transgenic flies were maintained at 25 °C and 70% humidity on standard corn meal and agar medium. The wing imaginal discs were obtained by

dissecting late third instar larvae and the expression patterns of EGFP and DsRed reporter genes were examined under a Leica DM 6000 fluorescent microscope.

Table 2.1 Sequences of primers and PCR products used for pCI-OMB construct cloning

PCR Product Name [a]	Primer Name	Sequence (5' to 3')	Comment [b]
lambda B (lambda DNA)	*NotI*-lambdaB lambdaB-*NheI*	ATTGCGGCCGCTGCTGCATGCGAAAGTCCTACGGTCAAGAG TTAGCTAGCAGATGCTGATATATTTTAGAGGTGA	*NotI* *NheI*
NheI-Insulator-*NheI* (Genomic DNA from flies *y v f mal flam*)	*NheI*-INSL *NheI*-RE-INSL	AACAGCTAGCGTTGTGTATCTGGCCACGTAA TGTTGCTAGCTTGGTTGTTGGTTGGCACAC	*NheI* *NheI*
FseI-Insulator-*KpnI* (Genomic DNA from flies *y v f mal flam*)	*FseI*-RE-INSL *KpnI*-INSL	TACAGGCCGGCCGTTGTGTATCTGGCCACGTAA GTAGGTACCTTGGTTGTTGGTTGGCACACC	*FseI* *KpnI*
NotI-Insulator-*NotI* (Genomic DNA from flies *y v f mal flam*)	MLJ-P23 MLJ-P24	TGTAGCGGCCGCGTTGTGTATCTGGCCACG TGTTGCGGCCGCTTGGTTGTTGGTTGGCAC	*NotI* *NotI*
ci enhancer L (Genomic DNA)	*ci* EN F new *ci* EN MR	CAACTAGTGTAACATGTCTAAGACCACAGGA CTCTTCTTTGGTTCAGTAC	*SpeI*
ci enhancer R (Genomic DNA)	*ci* EN MF *ci* EN R new2	CACAAGTGTGCGAGCATATC CGTTGACTAGTCTTGGCAATG	 *SpeI*
lambda A (lambda DNA)	Lambda A F Lambda A R	TTAGGTACCACCAGGAATGTAGTGGCGGA AATGCGGCCGCACCTGAGCTTAGAACCTTTAC	*KpnI* *NotI*
lambda C (lambda DNA)	Lambda C F Lambda C R	TTTCCTAGGCAGGACAACTTCTGGTCCGGT ATACCTAGGCTGTTCCATCGTGGTGATCCCGTT	*AvrII* *AvrII*
MluI OMB wing enhancer *SacII* (Genomic DNA)	*OMB* W LF *OMB* W RR	AAAGGTACCTTTACGCGTAGCAGCGTACTTCGTACTTCGT CAATCTAGAAATCCGCGGTGGCATAACCAAAACAAGAAATATGCCGA	*MluI* *SacII*
NotI-EGFP-*AscI* (pGreen H-Pelican)	*NotI* EGFP EGFP *AscI*	AATTGCGGCCGCGAGCGCCGGAGTATAAATAGA CGTGGCGCGCCATTAACGCTTACAATTTACGCCTT	*NotI* *AscI*
AvrII-DsRed-*BsiWI* (pRed H-Pelican)	*AvrII* DsRed DsRed *BsiWI*	TATCCTAGGAGCGCCGGAGTATAAATAGAG GTGCGTACGCATTAACGCTTACAATTTACGCC	*AvrII* *BsiWI*
pCR2.1-*SphI*-FRT-*AvrII*-DsRed-*BsiWI*(L) (pCR2.1-*AvrII*-DsRed-*BsiWI*)	*SphI*-F-*AvrII*-R DsRed *BsiWI*	TTCGCATGCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCTATCCTAGGAGCGCCGGAGT GTGCGTACGCATTAACGCTTACAATTTACGCC	*SphI* *BsiWI*
pCR2.1-*AvrII*-DsRed-*BsiWI*-FRT-*SphI* (R) (pCR2.1-*AvrII*-DsRed-*BsiWI*)	*AvrII*-DsRed *SphI*-F-*BsiWI*-R	TATCCTAGGAGCGCCGGAGTATAAATAGAG ATTGCATGCTGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGTGCGTACGCATTAACGCTTAC	*AvrII* *SphI*

[a] Names in parentheses indicate templates for PCR amplification.
[b] Underlined portions of the sequences represent restriction sites.

Table 2.2 Subcloning of DNA fragments into the pCR2.1 vector for the p*CI-OMB* constructs

Vector	Insert	Product
pCR2.1 (3929 bp)	lambdaB PCR product (1739 bp)	pCR2.1-lambdaB PCR (5668 bp)
pGreen H-Pelican (*Nhe*I/*Not*I: 10184bp)	pCR2.1-lambdaB PCR product (*NheI/NotI*: 1726 bp)	pGreen-H-Pelican with lambdaB (11910 bp)
pCR2.1 (3929 bp)	*NheI*-Insulator-*NheI* PCR product (464 bp)	pCR2.1-*NheI*-INSL-*NheI* (4393bp)
pCR2.1 (3929 bp)	*FseI*-Insulator-*KpnI* PCR product (465 bp)	pCR2.1-*FseI*-INSL-*KpnI* (4394 bp)
pCR2.1 (3929 bp)	*NotI*-Insulator-*NotI* PCR product (468 bp)	PCR2.1-*NotI*-INSL-*NotI* (4397 bp)
pCR2.1 (3929 bp)	*ci* enhancer L PCR product (1017bp)	pCR2.1-*ci* enhancer L (4946bp)
pCR2.1 (3929 bp)	*ci* enhancer R PCR product (1279bp)	pCR2.1-*ci* enhancer R (5208bp)
pCR2.1-*ci* enhancer L (*Kpn*I/*Age*I: 4854bp)	pCR2.1-*ci* enhancer R (KpnI/AgeI: 1266bp)	pCR2.1-*ci* enhancer (6120bp)
pCR2.1 (3929 bp)	lambdaA PCR product (1450 bp)	pCR2.1-lambdaA PCR (5379 bp)
pCR2.1 (3929 bp)	lambdaC PCR product (428 bp)	pCR2.1-lambdaC-PCR (4357 bp)
pCR2.1-*NotI*-loxP-INSL-loxP-*NotI* (*SacII/MluI*: 3980 bp)	*MluI-OMB* wing enhancer *SacII* PCR product (*MluI/SacII*: 2178 bp)	pCR2.1-*NotI*-loxP-*OMB*-loxP-*NotI* (6158 bp)
pCR2.1 (3929 bp)	*NotI*-EGFP-*AscI* PCR product (1215 bp)	pCR2.1- *NotI*-EGFP-*AscI* reverse (5144 bp)
pCR2.1 (3929 bp)	*AvrII*-DsRed-*BsiWI* PCR product (1165 bp)	pCR2.1- *AvrII*-DsRed-*BsiWI* (5094 bp)
pCR2.1 (3929 bp)	*SphI*-FRT-*AvrII*-DsRed-*BsiWI* (L) PCR product (1209 bp)	pCR2.1- *SphI*-FRT-*AvrII*-DsRed-*BsiWI* (L) (5138 bp)
pCR2.1 (3929 bp)	*AvrII*-DsRed-*BsiWI*-FRT-*SphI* (R) PCR product(1209 bp)	pCR2.1- *AvrII*-DsRed-*BsiWI*-FRT-*SphI* (R) (5138 bp)
pCR2.1- *SphI*-FRT-*AvrII*-DsRed-*BsiWI* (L) (*XhoI/AvrII*: 3938 bp)	pCR2.1- *AvrII*-DsRed-*BsiWI*-FRT-*SphI* (R) (*XhoI/AvrII*: 1244 bp)	pCR2.1-*SphI*-FRT-DsRed-FRT-*SphI* (5182 bp)
pCR2.1-*SphI*-FRT-DsRed-FRT-*SphI* (*AvrII*: 5182 bp)	pCR2.1-lambdaC-PCR (*AvrII*: 416 bp)	pCR2.1-*SphI*-FRT-lambdaC-DsRed-FRT-*SphI* (5598 bp)
pCR2.1-*NotI*-INSL-*NotI* (*NotI*: 3902 bp)	pCR2.1-*NotI*-INSL-*NotI* (*NotI*: 452 bp)	pCR2.1-*NotI*-INSL-*NotI* (*EcoRI* deletion) (4354bp)
pCR2.1-*NotI*-INSL-*NotI* (*EcoRI* deletion) (*EcoRI/SpeI*: 4329 bp)	pCR2.1-*ci*-enhancer (*SpeI*/*MfeI*: 2139 bp)	pCR2.1-*ci*-INSL (6468 bp)

Restriction sites, fragment sizes and product sizes are described in parentheses

Table 2.3 Cloning procedure for p*CI-OMB* constructs

	Vector	Insert	Product
A1	pCaSpeR4 (*EcoRI*: 7863 bp)	pCR2.1-*NotI*-EGFP-*AscI* reverse (*EcoRI*: 1231 bp)	pCaSper-EGFP (9094 bp)
A2	pCaSper-EGFP (*NotI*/*AgeI*: 8853 bp)	pGreen H-Pelican with lambda B (*NotI*/*AgeI*: 1932 bp)	pCaSpeR-lambda B-EGFP (10785 bp)
A3	pCaSpeR-lambda B-EGFP (*AvrII*: 10711 bp)	pCR2.1-*NheI*-INSL-*NheI* (*NheI*: 450 bp)	pCaSpeR-lambdaB-SG (11161 bp)
A4	pCaSpeR-lambdaB-SG (*SpeI*/*NotI*: 11147 bp)	pCR2.1-*FseI*-INSL-*KpnI* (*SpeI*/*NotI*: 533 bp)	pCaSpeR-S-lambdaB-SG (11680 bp)
A5	pCaSpeR-S-lambdaB-SG (*SpeI*: 11680 bp)	pCR2.1-*ci* enhancer PCR (*SpeI*: 2178 bp)	pCaSpeR-*ci*-S-lambdaB-SG (13858 bp)
A6	pCaSpeR-*ci*-S-lambdaB-SG (*KpnI*/*NotI*: 13813 bp)	pCR2.1-lambdaA PCR (*KpnI*/*NotI*: 1441 bp)	pCaSpeR-*ci*-S-lambdaAB-SG (15254 bp)
A7	pCaSpeR-*ci*-S-lambdaAB-SG (*NotI*: 15254 bp)	pCR2.1-*NotI*-loxP-*OMB*-loxP-*NotI* (*NotI*: 2257 bp)	pCaSpeR-*ci*-S-L*OMB*L-SG (17511 bp)
A8	pCaSpeR-*ci*-S-L*OMB*L-SG (*SphI*: 17511 bp)	pCR2.1-*SphI*-FRT-lambdaC-DsRed-FRT-*SphI* (*SphI*: 1657 bp)	**pCaSpeR-*ci*-S-L*OMB*L-FRF-SG (p*CI-OMB* RG)** **(19168 bp)**
B1	pCaSpeR-lambdaB-EGFP (*NheI*: 10785bp)	pCR2.1-*ci* enhancer PCR (*SpeI*: 2178 bp)	pCaSpeR-EGFP-*ci*-lambda B (12963 bp)
B2	pCaSpeR-EGFP-*ci*-lambda B (*SphI*: 11676bp)	pCR2.1-*SphI*-FRT-lambdaC-DsRed-FRT-*SphI* (*SphI*: 1657bp)	pCaSpeR-EGFP-*ci*-DsRed (13333 bp)
B3	pCaSpeR-EGFP-*ci*-DsRed (*AvrII*: 12824bp)	pCaSpeR-*ci*-S-L*OMB*L-SG (*SpeI*-*AvrII*: 4031 bp)	**pCaSpeR-EGFP-*ci*-S-*OMB*-DsRed (p*CI-OMB* 1)** **(16855 bp)**
C1	pCaSpeR4 (*XbaI* : 6468 bp)	pCR2.1-*ci*-INSL (*SpeI*/*XbaI* : 2639 bp)	pCaSpeR-*ci*-INSL (10488 bp)
C2	pCaSpeR-*ci*-INSL (*SphI*: 10488 bp)	pCR2.1-*SphI*-FRT-lambdaC-DsRed-FRT-*SphI* (*SphI*: 1657 bp)	pCaSpeR-*ci*-INSL-DsRed (12145 bp)
C3	pCaSpeR-*ci*-INSL-DsRed (*EcoRI*/*SacII*: 12115 bp)	pCaSpeR-*ci*-S-L*OMB*L-SG (*EcoRI*/*SacII*: 3372 bp)	pCaSpeR-EGFP-INSL-*ci*-INSL-DsRed (15487 bp)
C4	pCaSpeR-EGFP-INSL-*ci*-INSL-DsRed (*SpeI*/*SacII*: 15473 bp)	pCaSpeR-*ci*-S-L*OMB*L-SG (*SpeI*/*SacII*: 4142 bp)	**pCaSpeR-EGFP-S-*OMB*-S-*ci*-S-DsRed (p*CI-OMB* 2)** **(19615 bp)**

CHAPTER III

RESULTS

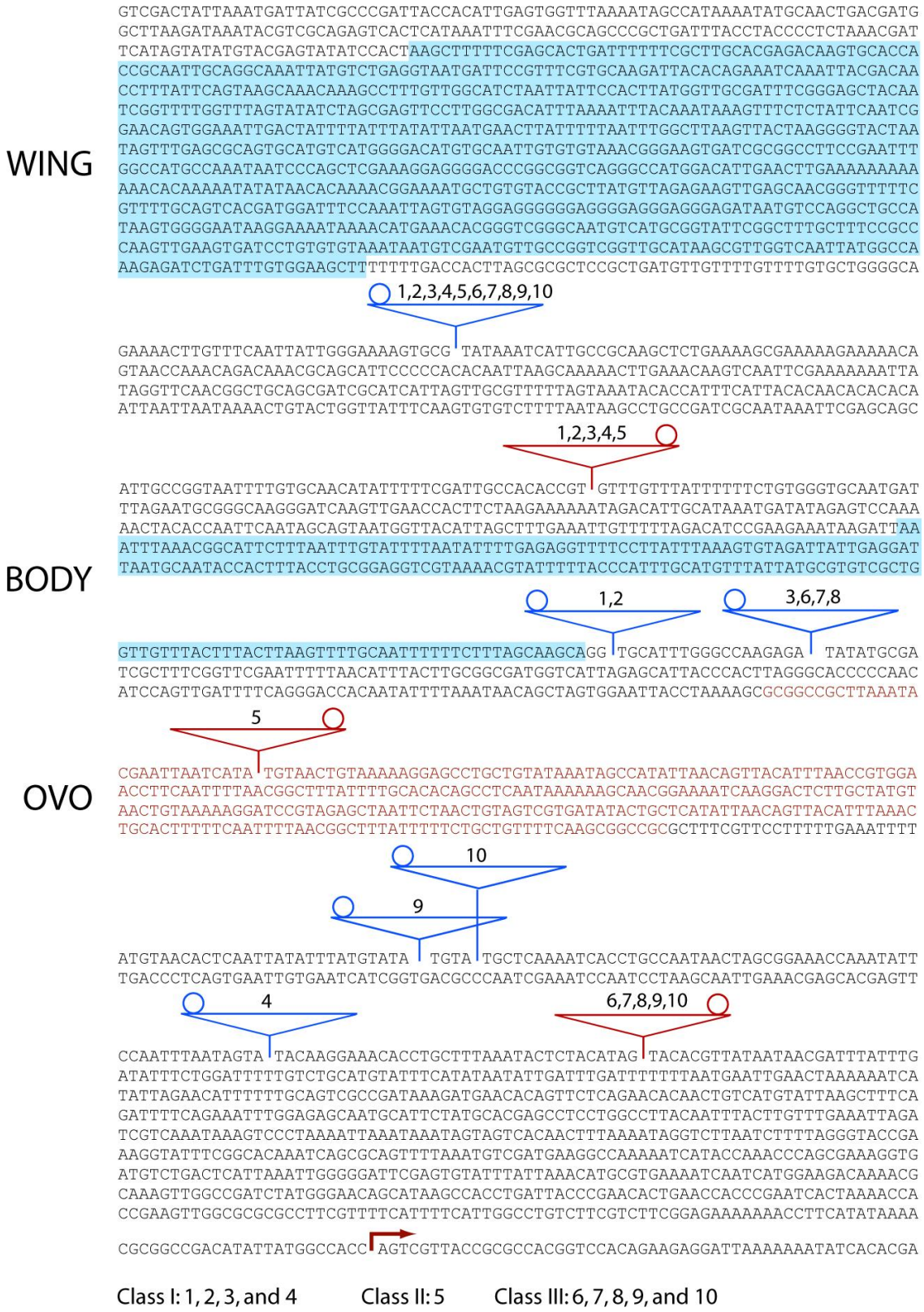
The relative orientation and position of three gypsy retrovirus insertions in the yellow transgene determine diverse patterns of enhancer-promoter interactions.

While pairing of two insulators has been shown to attenuate the enhancer blocking activity of insulators (Cai and Shen, 2001; Muravyova et al., 2001), the ability of three insulators to functionally interact to induce bypass of enhancer activity is unclear (Kuhn et al., 2003; Savitskaya et al., 2006). Therefore, to test whether three insulators induce bypass of enhancer activity and also whether the functional interactions of three insulators are dependent on the relative orientation of *gypsy* retroviruses, a third *gypsy* retrovirus was inserted into the regulatory region of the *yellow* gene containing two previously existing *gypsy* insertions. The third insertion was induced by crossing *flamenco* mutant lines with a transgenic *Casper-yellow* (*CasperY*) stock that carried two previously integrated *gypsy* elements in the regulatory region of the *yellow* transgene (Labrador et al., 2008). Two different stocks were used for the third insertion. One transgenic line carrying the *CasperY2.4y*$^{w+}$ transgene displays black wings and black body, and contains two *gypsy* insertions in opposite orientation integrated between the wing enhancer and body enhancer. The second stock used in this experiment (*CasperY2.4y*^{wB}) also has two *gypsy* insertions in opposite orientation, one between the wing and body enhancer and the second between the body enhancer

and the *yellow* promoter. *CasperY2.4y^{wB}* flies display black wings and yellow body. *CasperY2.4y^{w+}* and *CasperY2.4y^{wB}* were originally derived from the same initial *CasperY2.4*, into which three successive rounds of *gypsy* integrations were induced. Ten independent *Drosophila* lines carrying three *gypsy* insertions were identified by changes in the *yellow* phenotype after crossing with *flamenco* mutant females. Changes in the phenotypes of these lines cannot be due to position effects and were exclusively due to the activity of the *gypsy* insertions, given that the integration site of the initial *CasperY2.4* transgene is the same for all lines. Using the primers described in Methods we performed PCR combined with DNA sequencing to amplify and sequence all new *gypsy* integration sites (Figure 3.1). Results show that these ten lines could be grouped into 3 phenotypic classes, and that each phenotypic class corresponds to a specific arrangement of *gypsy* provirus orientations (Figure 3.2).

Phenotypic class I was defined by a lack of pigmentation in body cuticle and wing blades. This phenotype resulted from expression of the *CasperY2.4y^{w+2}* transgene, carrying three insulators arranged as in Figure 3.2A. A total of four lines were obtained with this phenotype, having two common *gypsy* insertions between wing and body enhancers and one independent insertion between the *yellow* promoter and body enhancer. In all four lines of class I, the third *gypsy* insertion is oriented in the same direction as the *yellow* gene. Regardless of the distance from the integration of the third insertion to the enhancers or to the promoter, these flies always show a *y^{2}* phenotype, suggesting that wing and body enhancers cannot communicate with the *yellow* promoter. Flies with black wings and yellow bodies were

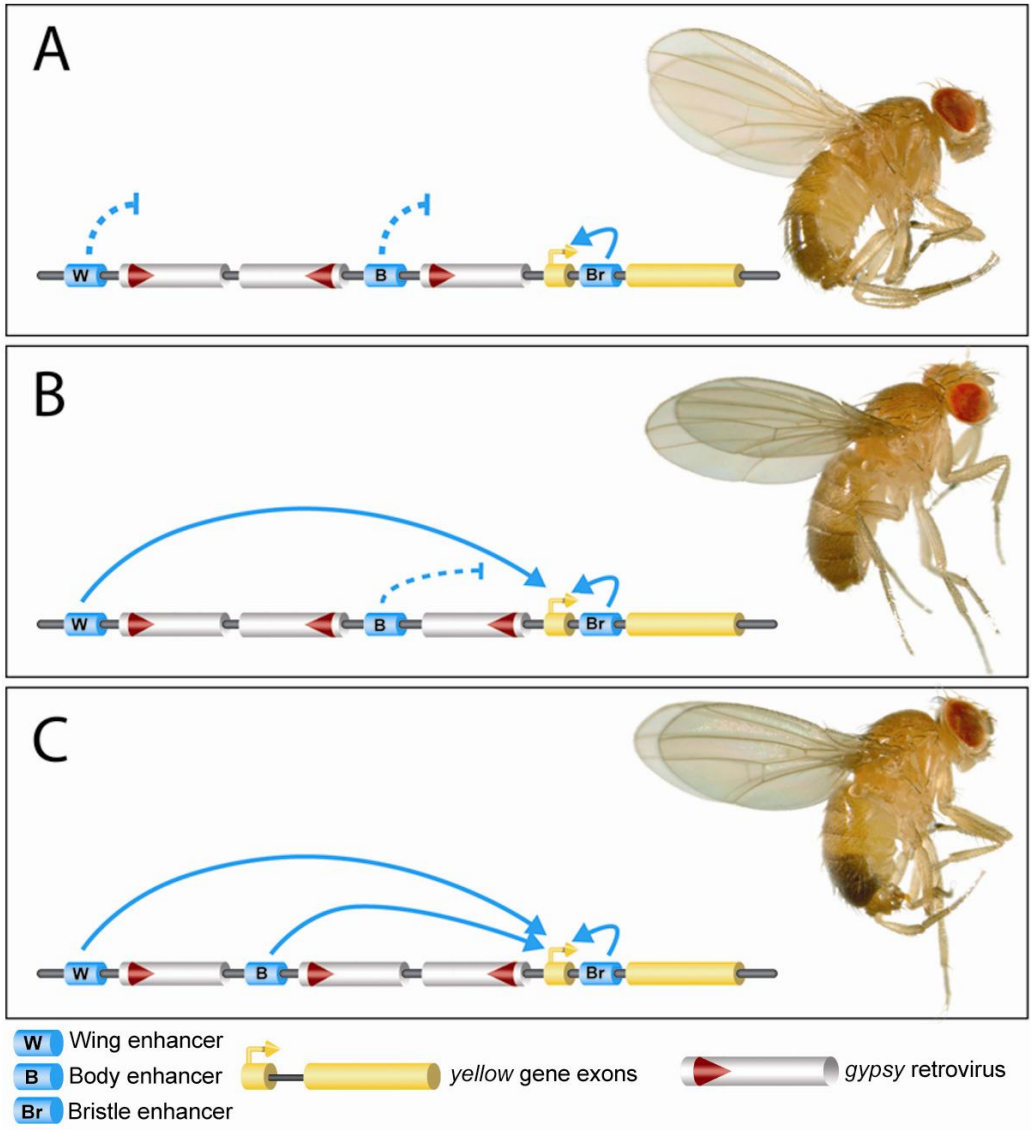
Figure 3.1 Map of *gypsy* retrovirus insertions into the regulatory region of the CaSpeR-*Yellow* (Casper*Y*) transgene. The transcription start site of the *yellow* transgene is indicated by the red arrow. Wing and body enhancer regions are highlighted in blue. Ovo binding sites, used for recruiting *gypsy* insertions into the *yellow* transgene (Labrador et al., 2008), are shown in red. The insertions and orientations of *gypsy* retroviruses are depicted as blue (5'-3') and red (3'-5') open triangles in the *yellow* transgene sequence. Open circles on open triangles indicate Su(Hw) binding regions in the *gypsy* retrovirus. Numbers indicate 10 independent transgenic lines grouped into three phenotypic classes: class I (1, 2, 3, and 4), class II (5), class III (6, 7, 8, 9, and 10). The numbers are marked on open triangles to describe the arrangement of three *gypsy* insertions in each of the transgenic lines.



Class I: 1, 2, 3, and 4 Class II: 5 Class III: 6, 7, 8, 9, and 10

Figure 3.2 Interactions between three *gypsy* insulators generate specific enhancer-promoter interactions that depend on the relative orientation of the *gypsy* insertions.

(A) Class I transgenic flies. Two *gypsy* elements are inserted in opposite orientation between wing and body enhancers. A third *gypsy* element is inserted with the 5' end toward the body enhancer. Transgenic flies show yellow wings and yellow body, indicating that the enhancer blocking activity of the *gypsy* insulators prevent the wing and body enhancers from activating the *yellow* promoter. **(B)** Class II transgenic flies. Insulators between the wing and body enhancers are identical to that of class I, but the promoter-proximal *gypsy* element is inserted in the opposite direction. Transgenic flies show black wings and yellow body. **(C)** Class III transgenic flies. Two *gypsy* insertions with opposite orientation are between the body enhancer and *yellow* promoter and a third *gypsy* insertion occurs between the wing and body enhancers, with the 5' end pointing toward the wing enhancer. Transgenic flies show black wings and black body, suggesting that wing and body enhancers bypass *gypsy* insulator activity thereby activating the *yellow* promoter.


A
W
B
Br
B
W
B
Br
C
W
B
Br
W Wing enhancer
B Body enhancer
Br Bristle enhancer
yellow gene exons
gypsy retrovirus

classified as phenotypic class II. Class II phenotypes result from the same *CasperY2.4y*$^{w+}$ transgene used to obtain class I, but with three *gypsy* insulators arranged in a different manner (*CasperY2.4y*$^{w+B}$; Figure 3.2B)**.** Only one line was included in class II, which has an arrangement of *gypsy* insertions identical to that of class I except for the proximal insertion, which in class II appears in opposite orientation to the *yellow* gene. Interestingly**,** although the body enhancer remains blocked, the change in orientation of the promoter-proximal *gypsy* insertion is sufficient to allow the wing enhancer to activate transcription at the *yellow* promoter. All five independent lines of phenotypic class III showed flies with black pigmentation both in body cuticle and in wing blades. Class III phenotypes result from transgene *CasperY2.4y*$^{wB+}$, which originates from the same initial *CasperY2.4* transgene as classes I and II but has three *gypsy* insulators arranged in a different manner (Figure 3.2C**).** The wild-type phenotype of these flies demonstrates that wing and body enhancers are capable of activating transcription at the promoter of the *yellow* transgene, regardless of the presence of three *gypsy* insertions, which add more than 23 kb of foreign DNA into the regulatory region of the transgene.

When the three phenotypic classes are compared with each other, only class III shows *yellow* gene expression activated by the body enhancer in the body cuticle. There are two *gypsy* insertions between the body enhancer and promoter in class III while a single *gypsy* insertion is present between the body enhancer and promoter in class I and II. This result is consistent with previous studies that two insulators placed between an enhancer and promoter attenuate the enhancer-blocking activity of insulator

(Cai and Shen, 2001; Muravyova et al., 2001). However, the possibility that enhancer-promoter communication may be affected by insulators located upstream of the enhancer, as well as insulators between the enhancer and promoter, cannot be excluded.

The effect of three insulator interactions can be observed clearly in wing tissue, since the three *gypsy* insertions are located between the wing enhancer and promoter. Class II and III transgenic flies show *yellow* gene activation by the wing enhancer. This result reveals that three insulators between the enhancer and the promoter can induce bypass of the enhancer and that the relative orientation of the middle *gypsy* retrovirus does not affect bypass of the wing enhancer. However, the different phenotypes in wing blades of class I and class II indicate that bypass of wing enhancer activity is sensitive to the relative orientation of the *gypsy* retrovirus located proximal to the promoter.

All three *gypsy* insulators participate in their interaction.

From the above data, it can reasonably be hypothesized that the two outer insulators preferentially interact and leave the middle one in an independent chromatin loop domain. Therefore, like the case of insertions of two *gypsy* retroviruses (Labrador et al., 2008), the relative orientations of the two outer retroviruses may determine bypass of enhancer activity but the presence or orientation of the middle *gypsy* insertion may not have any influence in the final output. To examine whether a third insulator located between two flanking insulators can influence insulator interaction, trangenes containing a GAL4-driven reporter gene with different arrangements of insulators were

constructed. Based on the observation that an increase in the distance between GAL4 binding sites and the promoter decreases the level of GAL4 activation in yeast (de Bruin et al., 2001; Guarente et al., 1984; Struhl, 1984), Mahmoudi et al used the GAL4-UAS system to show that the interaction between GAGA factors can mediate the activation of a promoter by the remote *cis*-regulatory region in yeast (Mahmoudi et al., 2002). Additionally, Kyrchanova et al showed interactions between diverse *Drosophila* insulators by using the UAS-GAL4 system (Kyrchanova et al., 2007). To generate the transgenes, we first introduced a 14X UAS and two reporter genes, EGFP and DsRed, into the pCaSpeR4 transformation vector and the construct was named P[G4ED] based on the order of arrangement (Figure 3.3A). When the P[G4ED] transgenic lines were crossed with a *GMR*-GAL4 transgenic line to activate the UAS-GAL4 driven reporter genes, the EGFP reporter gene proximal to the UAS site (20 bp) was expressed in eye tissue but the DsRed reporter gene distal to the UAS site (1.2kb) was not expressed (Figure 3.4A). This result is consistent with earlier observations that the GAL activator cannot stimulate the *white* promoter across the *yellow* gene in *Drosophila* (Kyrchanova et al., 2007).

To test whether interaction between two insulators can mediate the expression of the DsRed reporter gene driven by theUAS-GAL4 system, the P[G4(S)E(S)D] transgene, containing two insulators and lambda spacers, was prepared (Figure 3.3B). Insertions of lambda DNA spacers and insulators increased the distance between the GAL4 binding sites and the two reporter genes. The distances of the EGFP promoter and the DsRed promoter from the UAS site were 2.3 kb and 5.0 kb, respectively. One

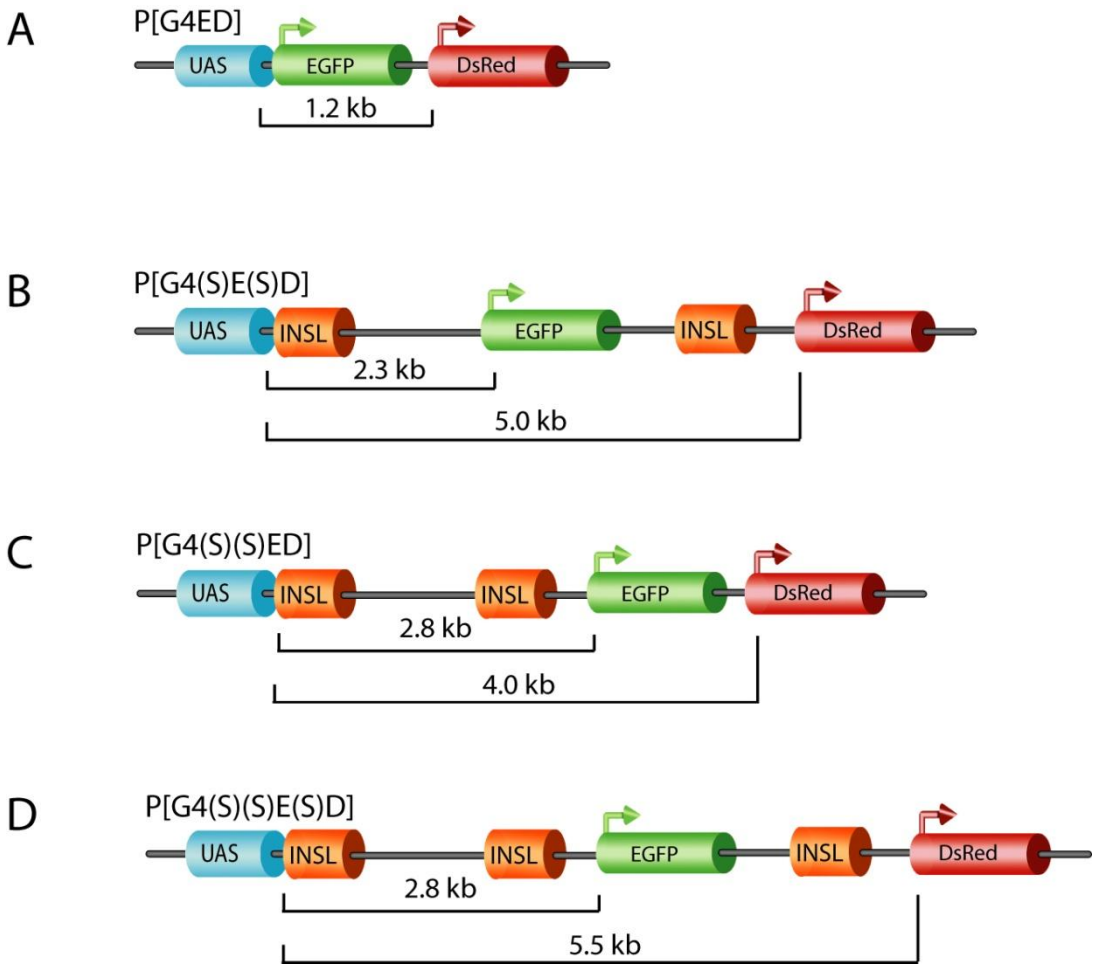
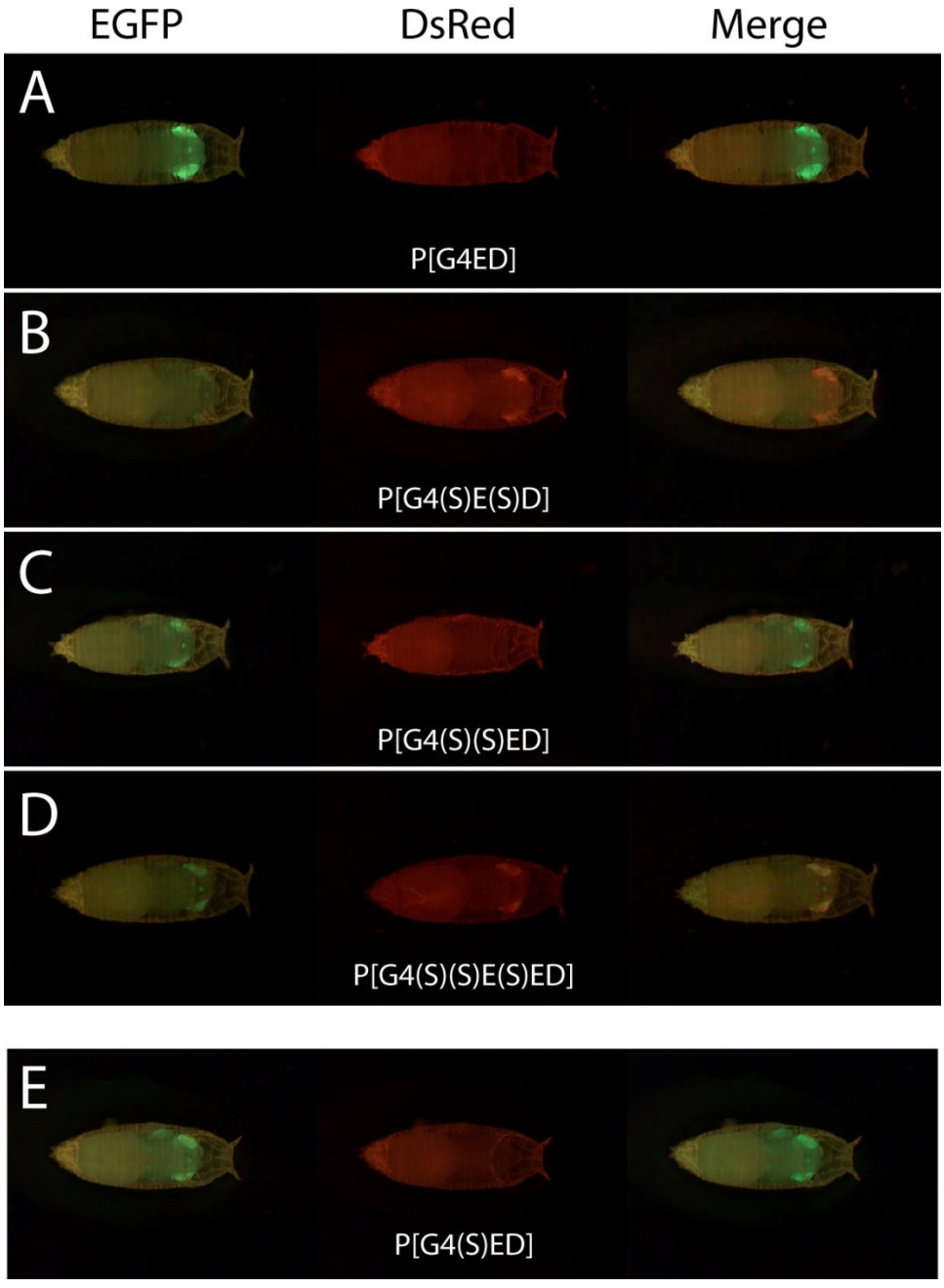


Figure 3.3 Schematic drawing of Gal4-UAS driven reporter transgenes.

Two reporter genes, EGFP and DsRed, are illustrated as green and red cylinders, respectively. The blue cylinder indicates the 14X GAL4 binding sites, and *gypsy* insulator sequences are depicted as orange cylinders. The names of the transgenes and the distance between the UAS sites and reporter genes are shown on the top and the bottom of each diagram, respectively.

Figure 3.4 The eye-specific expression pattern of the reporter gene driven by *GMR*-Gal4 in pupal stage of transgenic lines.

(A) Strong expression of the EGFP reporter gene is induced, whereas no DsRed expression is shown in P[G4ED]. **(B)** In P[G4(S)E(S)D] lines, the DsRed reporter gene is strongly activated and the expression level of EGFP is significantly decreased, showing red eyes in the merged image. **(C)** The Gal4-UAS system induces the expression of the EGFP reporter gene, but fails to activate the DsRed reporter gene in P[G4(S)(S)ED] lines. Compared with P[G4ED], P[G4(S)(S)ED] shows a strong, but decreased, level of EGFP expression. **(D)** In P[G4(S)(S)E(S)D], EGFP and DsRed reporter genes are activated with a similar expression level, showing yellow eyes in the merged image. **(E)** In P[G4(S)ED] lines, the Gal4-UAS system induces strong expression of the EGFP reporter gene in spite of the existence of the insulator between UAS and EGFP.


EGFP
DsRed
Merge
A
P[G4ED]
B
P[G4(S)E(S)D]
C
P[G4(S)(S)ED]
D
P[G4(S)(S)E(S)ED]
E
P[G4(S)ED]

gypsy insulator was added next to the GAL4 binding sites and the other insulator was placed approximately 400 bp upstream of the DsRed promoter. In spite of the 5.0 kb distance between the UAS and DsRed promoter, P[G4(S)E(S)D] lines showed strong expression of DsRed (Figure 3.4B). Considering that the GAL4 activator failed to activate the DsRed reporter gene located 1.2kb from the 14X UAS site in the P[G4ED] transgenic lines, the expression of the DsRed reporter gene in the P[G4(S)E(S)D] lines suggests that the interaction of two insulators led to the access of the DsRed promoter by the GAL4 binding sites. However, GAL4 activation also resulted in the expression of the EGFP reporter gene in the P[G4(S)E(S)D] transgenic lines, although the level of EGFP expression was significantly less than that in the P[G4ED] transgenic lines. Even if *gypsy* insulators are known to be able to block more than 20 different enhancers (Brasset and Vaury, 2005), selective enhancer-blocking has been shown for some specific promoters or enhancers (Cai and Levine, 1995; Muller, 2000; Scott et al., 1999). When we tested a transgene in which one insulator was added between UAS and EGFP reporter gene of P[G4ED] (P[G4(S)ED]), we could still observe the strong expression of EGFP in the P[G4(S)ED] lines (Figure 3.4E).

To confirm whether the expression of the DsRed reporter gene in P[G4(S)E(S)D] was due to the interaction between the insulator next to the 14X UAS and the other insulator upstream of the DsRed promoter, the P[G4(S)(S)ED] transgene was tested by crossing with *GMR*-GAL4 lines. Like P[G4(S)E(S)D], P[G4(S)(S)ED] contains two insulators. One insulator was placed next to the GAL4 binding sites and the other insulator was placed approximately 600 bp upstream of the EGFP promoter. The

distances of the EGFP promoter and DsRed promoter from the UAS site in P[G4(S)(S)ED] were 2.8 kb and 4 kb, respectively (Figure 3.3C). Even though the distance between the 14X UAS and DsRed promoter in P[G4(S)(S)ED] was approximately 1kb shorter than the distance in P[G4(S)E(S)D], expression of the DsRed reporter gene was not induced (Figure 3.4C). This result indicates that the strong expression of DsRed in P[G4(S)E(S)D] was due to the insulator located upstream of the DsRed promoter. In addition, the level of EGFP expression in P[G4(S)(S)ED] was stronger than that in P[G4(S)E(S)D], suggesting that the UAS-GAL4 driven EGFP activation was induced by the interaction between the insulator adjacent to the UAS site and the insulator upstream of the EGFP promoter.

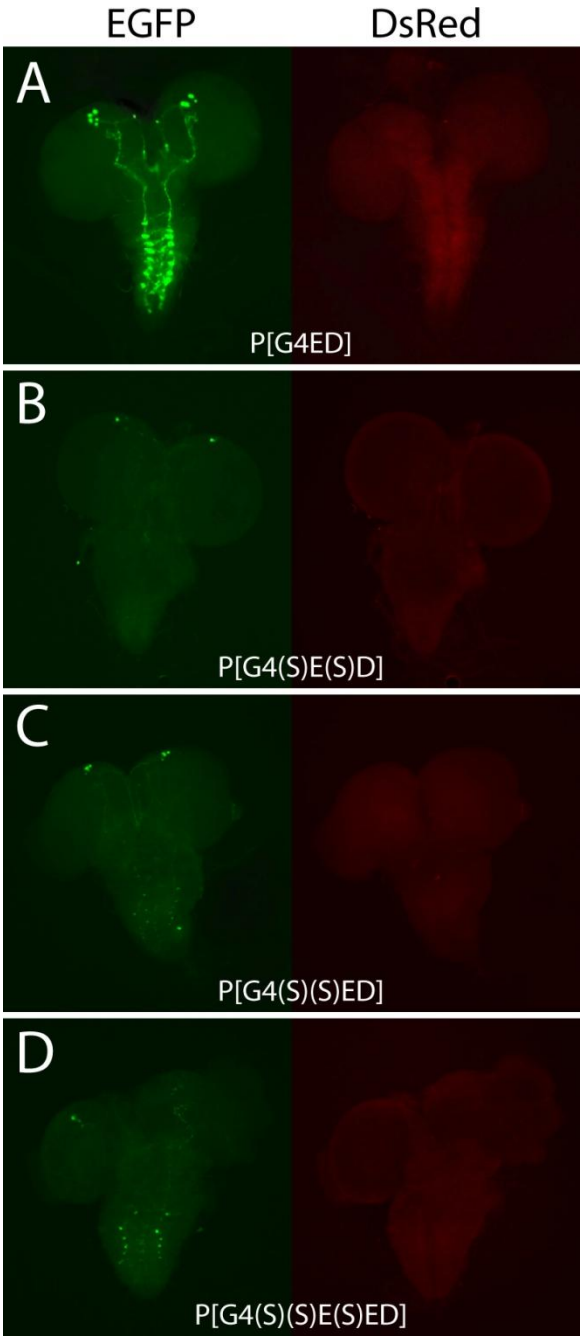
Next, the effect of three insulators on the expression of reporter genes was examined by crossing P[G4(S)(S)E(S)D] transgenic flies with *GMR*-GAL4 lines. Like in previous transgenes, the P[G4(S)(S)E(S)D] transgene has one insulator next to 14X UAS. The second and third insulators were placed at positions identical with the insulator upstream of the EGFP promoter in P[G4(S)(S)ED] and the insulator upstream of the DsRed promoter in P[G4(S)E(S)D], respectively (Figure 3.3D). Therefore, the distance of the EGFP promoter from the UAS site is 2.8 kb and the distance between the DsRed promoter and the UAS is 5.5 kb. Interestingly, the P[G4(S)(S)E(S)D] transgenic lines showed a specific expression pattern that is distinct from that of P[G4(S)E(S)D] and P[G4(S)(S)ED] containing two *gypsy* insulators. Both EGFP and DsRed reporter genes are expressed with equal intensities in eye tissues (Figure 3.4D). This result suggests that all three insulators participate in insulator interaction. If only

the two outer insulators interact without involvement of the middle insulator adjacent to EGFP, strong expression of DsRed and weak expression of EGFP should be induced, as seen in P[G4(S)E(S)D] lines. Interestingly, the expression of the EGFP reporter gene in a loop domain was driven by GAL4-UAS located outside of the loop domain, showing that the formation of loop domains by three insulator interactions could not block GAL4-UAS driven reporter gene expression. The specific phenotype of the P[G4(S)(S)E(S)D] lines, which is distinct from transgenic lines containing two *gypsy* insulators, may result from interactions among all three insulators or, alternatively, only two of the three insulators may interact without any preference of interaction among them. In transgenes containing two insulators, the insulator next to the UAS site may have no choice but to interact with the other insulator upstream of one reporter gene in all eye cells, while, in P[G4(S)(S)E(S)D], the insulator next to the UAS site may interact with the insulator upstream of the EGFP promoter in a fraction of eye cells, and with the insulator upstream of the DsRed promoter in other cells. This may appear as if both DsRed and EGFP reporter genes are being simultaneously expressed in the eye tissue of P[G4(S)(S)E(S)D].

To test whether EGFP and DsRed reporter genes are simultaneously expressed in a single cell, our transgenic lines were crossed with *crz*-GAL4 lines. Corazonin (*Crz*), is an insect neuropeptides that is expressed in third instar larvae in only 24 neuron cells: one pair of dorsomedial neurons, three pairs of dorsolateral neurons, and eight pairs of ventral nerve cord neurons (Choi et al., 2005). The expression of reporter genes at the single-cell level was investigated by dissecting the central nervous system of third instar

Figure 3.5 The expression pattern of the reporter gene driven by *Crz*-Gal4 in the central nervous system (CNS) of transgenic third instar larvae.

(A) In P[G4ED] lines, strong EGFP expression is induced, but DsRed expression is not induced. The similar expression pattern of reporter genes are shown in P[G4ED] lines crossed with *GMR*-Gal4 lines (Figure R3A). **(B, C, and D)** In all other transgenic lines (P[G4(S)E(S)D, P[G4(S)(S)ED], and P[G4(S)(S)E(S)D]) containing *gypsy* insulator sequences and a lambda spacer, the DsRed reporter genes are not expressed in neuron cells at all and weak expression of EGFP reporter genes is detected, showing the constant EGFP signal only in the neuron cells of the dorsolateral region.


EGFP
DsRed
A
P[G4ED]
B
P[G4(S)E(S)D]
C
P[G4(S)(S)ED]
D
P[G4(S)(S)E(S)ED]

larvae (Figure 3.5). Unexpectedly, crossing with C*rz*-Gal4 caused expression patterns different from crossing with *GMR*-Gal4. Similar to crossing with *GMR*-GAL4, crossing of P[G4ED] with *Crz*-Gal4 caused strong expression of EGFP and no expression of DsRed. However, all other transgenic lines showed only weak EGFP expression in neuron cells. This result suggests that insulator proteins or other factors crucial for interaction between insulators may not be expressed in CNS neuron cells, and so the GAL4-UAS regulatory element may not be able to be placed close to the promoters of reporter genes.

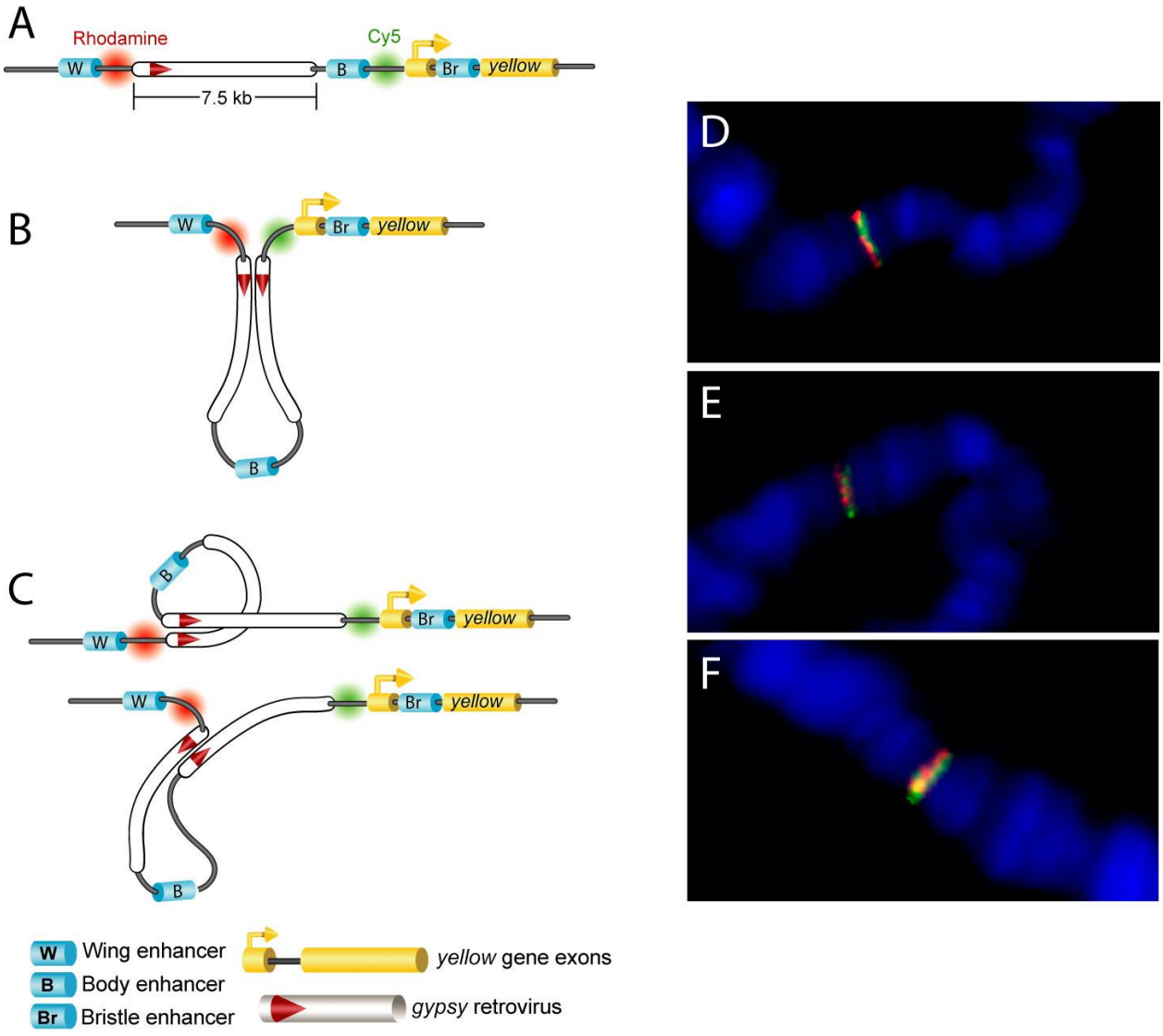
Gypsy retrovirus may contain repressive factor(s), in addition to insulators, that affect enhancer-promoter communication.

In all our transgenic flies containing insertions of two or three *gypsy* retroviruses, the wing enhancer activates the *yellow* promoter in transgenic lines which have a 3' to 5' insertion of *gypsy* retrovirus proximal to the promoter (Figure 1.5B, 3.2B, and 3.2C), whereas the wing enhancer cannot activate the promoter when the *gypsy* retrovirus proximal to *yellow* promoter is placed in the 5' to 3' direction (Figure 1.5C and Figure 3.2A). The relative orientation of the *gypsy* insertion proximal to the promoter leads the insulator region to be placed close to the promoter in the former and far away from the promoter in the latter. Considering that the ability of the enhancer to bypass two insulators is due to the interaction between two insulators, which allows enhancers to access promoters, the failure to activate the *yellow* gene by the wing enhancer in the

latter indicates that wing enhancer activity is affected by the remaining ~7 kb *gypsy* DNA between promoter and the wing enhancer. A possible explanation is that the ~7 kb intervening *gypsy* DNA may generate a distance too great for the wing enhancer to reach the promoter, such that even when the two elements interact, the enhancer and the promoter are unable to physically interact with one another as a result of this intervening DNA. To investigate whether the effect of the distance according to the relative orientation of *gypsy* insertions can be detected on polytene chromosomes, the transgenic line with a single *gypsy* insertion between the wing enhancer and body enhancer (Figure 1.3C), as well as transgenic lines of double insertions with different relative orientations of the second *gypsy* element inserted between body enhancer and promoter (Figure 1.4B and 1.4C), were used for fluorescence *in situ* hybridization (FISH). One DNA sequence probe was obtained by PCR amplifying a 0.5 kb sequence region close to the wing enhancer, and the other probe was obtained by amplifying a 0.5 kb upstream region close to the promoter. A single *gypsy* element insertion of 7.4kb would increase the distance between the two probes on the *yellow* transgene (Figure 3.6.A). If two insulators interact with each other on polytene chromosomes, the interaction of the two insulators in transgenic lines with *gypsy* insertions in opposite orientations will decrease the distance between the two probes and so result in overlapping of the two probe signals (Figure 3.6B). In transgenic lines with two *gypsy* insertions in the same orientation, the insulator interaction will still leave approximately 7kb distance between the two probes (Figure 3.6C). However, the three different transgenic lines tested did not show any significant differences in the band patterns of

Figure 3.6 Cytological analysis of insulator interactions on polytene chromosomes by fluorescence *in situ* hybridization.

Rhodamine and Cy5 were used for the probes close to the wing enhancer and close to the promoter, respectively. The rhodamine signal is shown in red and Cy5 signal is shown in green. DNA stained with DAPI is shown in blue. **(A)** In the *yellow* transgene containing a single *gypsy* insertion, the approximately 7.5 kb *gypsy* element would separate the two probe signals on polytene chromosomes. **(B)** When two insulators in the opposite orientation interact, the two probe signals would be close to each other. **(C)** The diagrams show two possibilities of insulator interaction in the transgene with two *gypsy* retrovirus insertions in the same orientation. In either of the two possibilities, the two probe signals would be separated by a distance of approximately 7 kb. **(D), (E),** and **(F)** show the actual FISH signal patterns on the polytene chromosomes of transgenic lines described in (A), (B), and (C), respectively.


A
Rhodamine
Cy5
W
B
Br
yellow
7.5 kb
B
W
Br
yellow
C
B
W
Br
yellow
W
Br
yellow
B
D
E
F
W Wing enhancer
B Body enhancer
Br Bristle enhancer
yellow gene exons
gypsy retrovirus

the two probes on polytene chromosomes (Figure 3.6D, E, and F). In our recent study of the distribution of Su(Hw) proteins on polytene chromosomes using fluorescence *in situ* hybridization (FISH) and immunostaining, independent endogenous Su(Hw) binding sites at distances as low as 10kb could be shown either as clearly separated sites or as a single band, depending on the local chromatin structure (Wallace et al., 2009), indicating that the level of resolution of FISH and immunostaining with polytene chromosomes is limited to distances >10 kb. Hence, higher resolution techniques will be needed to dissect these interactions further. As an approach to understanding the phenotypic changes dependent on the relative orientation of the *gypsy* insertion proximal to the promoter, *yellow* phenotypic recovery of transgenic lines carrying two *gypsy* retrovirus insertions were examined in homozygous null *su(Hw)* mutants. In the *su(Hw)* mutant background, body pigmentation was considerably rescued, revealing that the Su(Hw) binding sites in the *gypsy* retrovirus do not function as insulators in the absence of the Su(Hw) protein. Wing pigmentation in transgenic lines with two *gypsy* insertions in opposite orientation is decreased in the *su(Hw)* mutant background (compare Figure 3.7A with Figure 3.7B), whereas transgenic lines with the same orientation restored wing pigmentation in the absence of the Su(Hw) protein (compare Figure 3.7C and Figure 3.7D). Interestingly, as a consequence of these phenotypic changes, the resulting levels of *yellow* gene expression in wing blades were very similar in the *su(Hw)*$^{-}$ flies with two *gypsy* insertions, regardless of their relative orientation. These results indicate that access of the wing enhancer to the promoter for activation is controlled by both the insulator interaction and the relative orientation of the *gypsy*

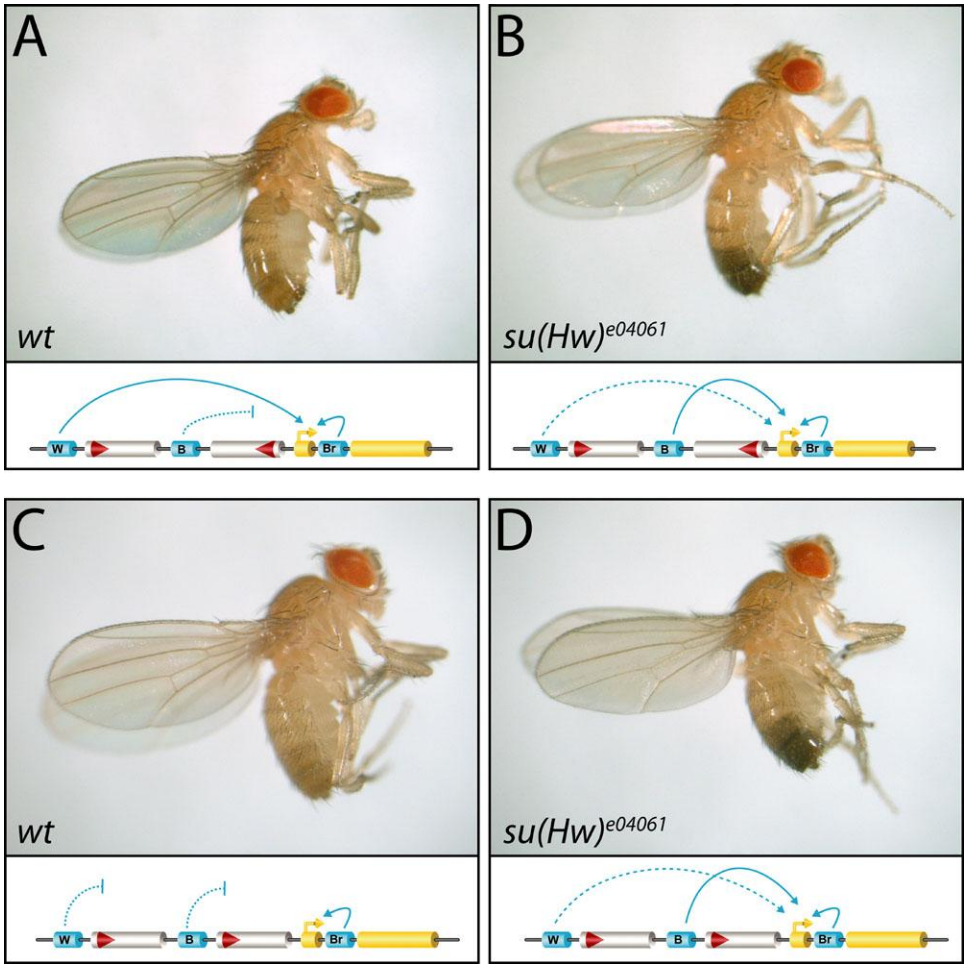


Figure 3.7 Phenotypic changes of transgenic lines with two *gypsy* insertions in the *su(Hw)*^{e04061} mutant background.

(A) Transgenic flies containing two *gypsy* insertions in the opposite orientation show a black wing and yellow body phenotype. **(B)** In the *su(Hw)* mutant background in (A), the body pigmentation is recovered from yellow to black, but wing pigmentation decreases to a level less than wild-type. **(C)** Transgenic flies containing two *gypsy* insertions in the same orientation show a yellow wing and yellow body phenotype. **(D)** In the *su(Hw)* mutant background in (C), both the body and wing pigmentations are increased but the degree of wing pigmentation recovery is very weak.

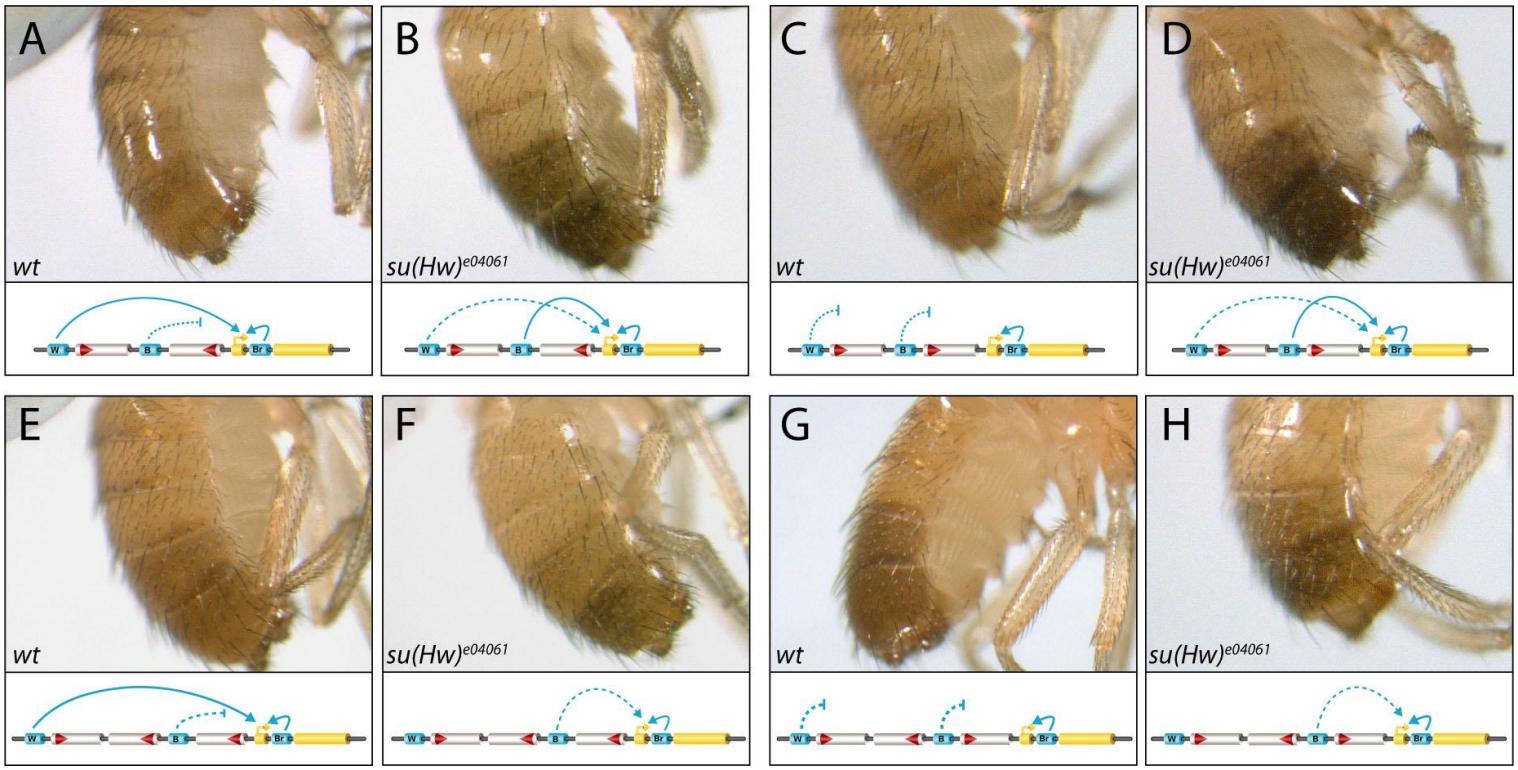
insertions.

Pigmentation of wings and body cuticle in the y^2 allele caused by insertion of a single *gypsy* retrovirus at -700 bp from the transcription start site of the endogenous *yellow* gene are rescued from levels of the null *yellow* allele to wild-type levels in the *su(Hw)* mutant background (Gerasimova et al., 1995; Ramos et al., 2006). However, in our transgenic flies containing two *gypsy* insertions, body pigmentation was rescued to almost wild-type levels, but wing pigmentation was very weak compared with that of wild-type flies. This may be due to the increase of the relative distance between wing enhancer and promoter by integration of two *gypsy* retroviruses. To test whether the weak *yellow* expression in wing tissue resulted from the two-fold increase in distance between the wing enhancer and the promoter by two *gypsy* insertions, the effect of a *su(Hw)* null mutation in transgenic lines with three *gypsy* insertions (class I and class II) was tested. The $su(Hw)^-$ flies failed to restore pigmentation in wing tissue, suggesting that an additional insertion of a third *gypsy* retrovirus between the wing enhancer and promoter may lead to a distance that is too great for the wing enhancer activity to reach to the *yellow* promoter.

Compared with the recovery of body pigmentation in $su(Hw)^-$ flies carrying two *gypsy* insertions (Figure 3.8B and D), pigmentation in the body cuticles of $su(Hw)^-$ flies carrying three *gypsy* insertions was not significantly rescued (Figure 3.8F and H). This result was surprising, because there is no difference in distance between the body enhancer and promoter in both two and three *gypsy* insertion transgenic lines. This finding suggests that the decrease of wing enhancer activity according to the increase in

Figure 3.8 Comparison of body pigmentation of transgenic lines with two or three *gypsy* insertions in the *su(Hw)*^{e04061} mutant background

(A-D) Mutation of Su(Hw) in transgenic lines with two *gypsy* insertions results in the considerable recovery of body pigmentation. **(E-H)** The body pigmentation of transgenic lines with three *gypsy* insertions is not significantly recovered in the *su(Hw)* mutant background. When the patterns of the *gypsy* insertion between the body enhancer and the promoter are compared, F and H are identical to B and D, respectively. Nevertheless, the level of body pigmentation in *su(Hw)* mutant flies with three *gypsy* insertions is significantly decreased.

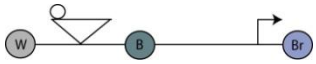
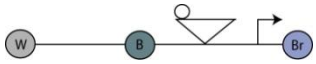
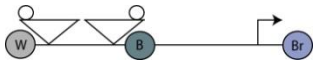
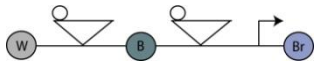
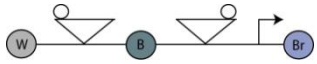
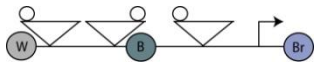
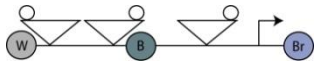
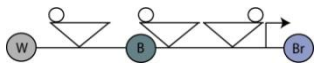

A
wt
B
su(Hw)e04061
C
wt
D
su(Hw)e04061
E
wt
F
su(Hw)e04061
G
wt
H
su(Hw)e04061

number of *gypsy* insertions in the *su(Hw)* mutant background may result at least partially from other factors involved in the repression of gene expression that may exist in the *gypsy* retrovirus.

The absence of Mod(mdg4)67.2 protein results in local gene silencing.

It has been reported that Modifier of mdg4 67.2 [Mod(mdg4)67.2] interacts with itself as well as with Suppressor of Hairy wing [Su(Hw)], suggesting that interactions between two insulators are mediated by interactions between Mod(mdg4)67.2 proteins (Ghosh et al., 2001; Pai et al., 2004). Therefore, the lack of *mod(mdg4)67.2* might result in the removal of insulator interactions. To investigate the effect of mutations in the *mod(mdg4)* gene on phenotypes induced by *gypsy* insertions in the *yellow* transgenes**,** transgenic lines carrying one, two, or three *gypsy* insertions were established as homozygous $mod(mdg4)^{u1}$ alleles. The $mod(mdg4)^{u1}$ mutant allele, caused by the insertion of the *Stalker* retrotransposon into the sequences encoding the carboxyl-terminal end of the Mod(mdg4)67.2 protein, is a hypomorph that expresses low levels of a truncated form of the Mod(mdg4)67.2 protein (Gerasimova et al., 1995). All transgenic lines, except for two lines, showed yellow wings, yellow body, and variegated bristles in a $mod(mdg4)^{u1}$ mutant background (Table 3.1). Considering that the bristles enhancer and the *yellow* promoter are located downstream of the insulator(s) and the enhancer-blocking activity of the insulator is unidirectional, the variegated bristle phenotype suggests that the absence of Mod(mdg4)67.2 protein causes the repression of transcription at the *yellow* promoter. Therefore, it could be argued that the yellow

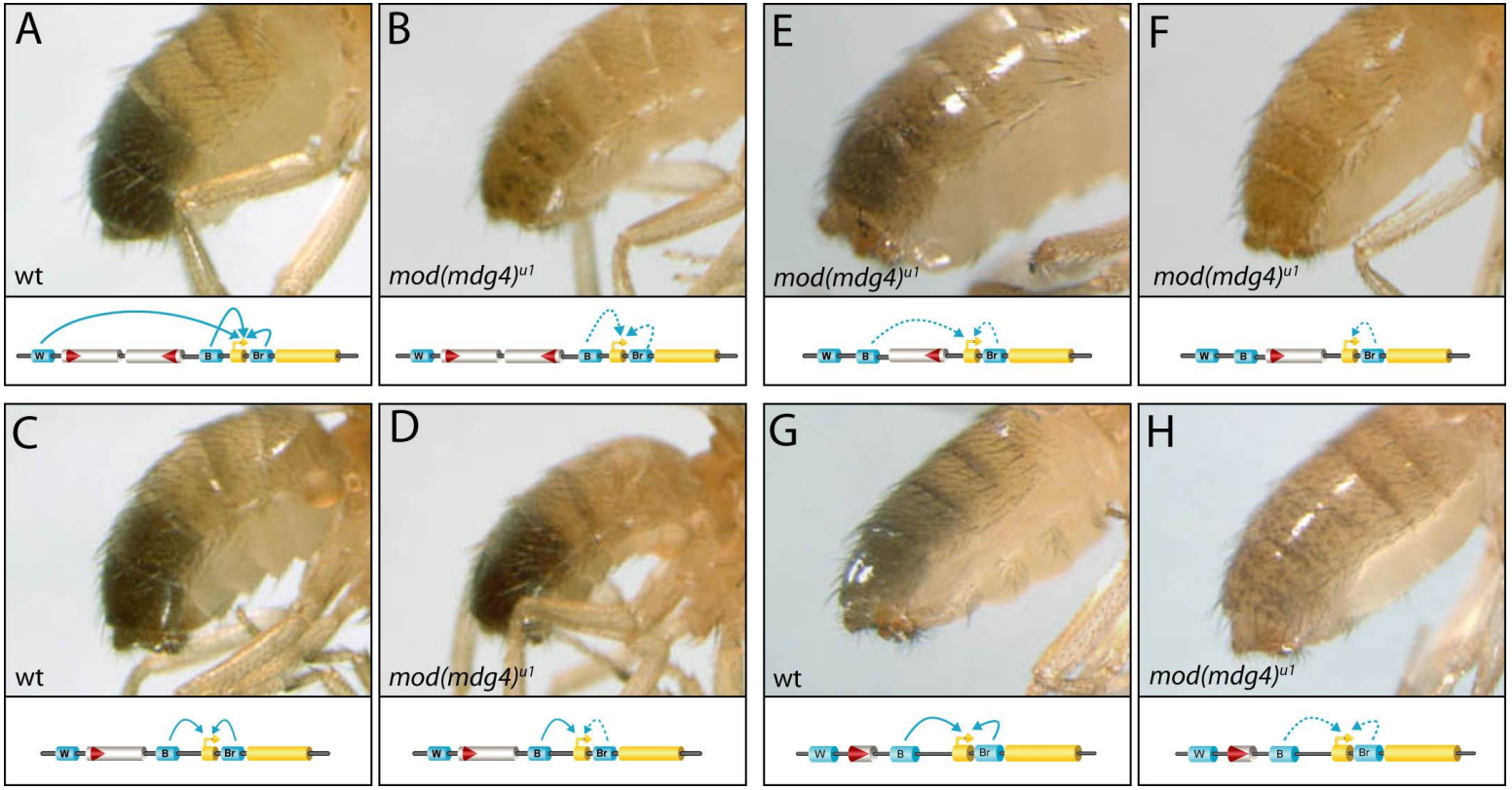
Table 3.1 Effects of *mod(mdg4)* mutation on the *yellow* phenotypes of transgenic lines with *gypsy* insertions

Strain		yellow phenotype		
		Wing	Abdomen	Bristles
Oregon R		+++	+++	+++
yw^{67C}		-	-	-
	wt	-	+++	+++
	mod(mdg4)^{u1}	-	+++	variegation
	wt	-	-	+++
	mod(mdg4)^{u1}	-	-	variegation
	wt	+++	+++	+++
	mod(mdg4)^{u1}	-	variegation	variegation
	wt	-	-	+++
	mod(mdg4)^{u1}	-	-	variegation
	wt	+++	-	+++
	mod(mdg4)^{u1}	-	-	variegation
	wt	-	-	+++
	mod(mdg4)^{u1}	-	-	variegation
	wt	++	-	+++
	mod(mdg4)^{u1}	-	-	variegation
	wt	++	++	+++
	mod(mdg4)^{u1}	-	-	variegation

Wing, body and bristle enhancers are described as colored circles with their initials. Arrow bar indicates promoter. The insertions of *gypsy* retroviruses in the *yellow* transgene are described as triangles on the line and the circles on triangles indicate insulator region in *gypsy* retrovirus. Phenotypes were quantitated by visual inspection under a dissecting microscope. Levels of coloration were assigned values between null (minus) and wild type (triple plus).

Figure 3.9 The effects of the *mod(mdg4)^{u1}* mutation

(A) Two *gypsy* retroviruses are placed upstream of body enhancer and so body enhancer activates the promoter in wild type, showing black pigmentation in body cuticle. **(B)** The impaired black pigmentation in body tissue is shown as dots in *mod(mdg4)^{u1}* mutant of (A). **(C)** The insertion of single *gypsy* retrovirus upstream of body enhancer induces the strong expression of *yellow* gene in the body cuticle. **(D)** The *mod(mdg4)^{u1}* mutation does not result in any change of body cuticle pigmentation of (C). **(E)** the *y^{2}* allele with *gypsy* retrovirus insertion between body enhancer and the promoter in endogenous *yellow* gene shows variegated body pigmentation in *mod(mdg4)^{u1}* mutant background. **(F)** In transgenic lines that contain *gypsy* insertion pattern identical to the *y^{2}* allele except for the orientation of insertion, the *mod(mdg4)^{u1}* mutation does not cause variegated body pigmentation. **(G)** The transgenic lines containing a single insulator upstream of the body enhancer show black body pigmentation, like the transgenic lines with a single *gypsy* retrovirus upstream of the body enhancer (Compare C and G). **(H)** The variegated body pigmentation is induced in the *mod(mdg4)^{u1}* mutant of (G), while there is no change in the *mod(mdg4)^{u1}* mutant line containing a single *gypsy* retrovirus upstream of the body enhancer (Compare D with H).


A
wt
B
mod(mdg4)u1
E
mod(mdg4)u1
F
mod(mdg4)u1
C
wt
D
mod(mdg4)u1
G
wt
H
mod(mdg4)u1

coloration in wing and body tissues may be due to gene silencing. One transgenic line, which contains two *gypsy* insertions between wing enhancer and body enhancer and shows a wild-type phenotype, showed a variegated pattern in coloration of body cuticle in the *mod(mdg4)*^{u1} mutant background (Figure 3.9B). Like the bristles enhancer and promoter described above, the body enhancer and promoter of this transgenic line are located downstream of the insulators. The variegated body phenotype, which is contradictory to the unidirectional insulator function, supports the idea that gene silencing is induced by the lack of Mod(mdg4)67.2 protein. A second line showed a phenotype identical to wild-type in the *mod(mdg4)*^{u1} mutant background with the exception of variegated bristles (Figure 3.9D). This transgenic line contains an insertion of a single *gypsy* retrovirus between wing enhancer and body enhancer and shows a yellow wing and a black body phenotype in the presence of the Mod(mdg4)67.2 protein. Because there is no position effect in all the transgenic lines containing *gypsy* insertions, it was surprising that the body enhancer of this transgenic line was not affected by gene silencing in *mod(mdg4)*^{u1} mutant background. It has been already reported elsewhere that gene silencing is induced in the absence of the Mod(mdg4) protein (Gerasimova et al., 1995). The y^2 allele, caused by insertion of a single *gypsy* retrovirus between the body enhancer and the promoter in the endogenous *yellow* gene, showed variegated body cuticle and bristle phenotype in homozygous *mod(mdg4)*^{u1} alleles (Figure 3.9E). It was suggested from this result that the absence of Mod(mdg4) results in free leucine zipper and acidic domains of Su(Hw), and these free domains may interact with other proteins perhaps involved in the condensation or

formation of heterochromatin (Gerasimova et al., 1995). Interestingly, our *mod(mdg4)*^{u1} mutant line carrying a *single gypsy* insertion between the body enhancer and the promoter did not show any variegated body phenotype (Figure 3.9F). This may be due to the position effect caused by the difference between the endogenous *yellow* gene and ectopic *yellow* transgene, but another possibility is that the difference in orientation of the single *gypsy* insertion between the y^2 allele and our transgenic line may cause the different *yellow* phenotype in body tissue. While the insulator sequences in the *gypsy* element are placed close to the promoter in the y^2 allele, our transgenic line has a single *gypsy* insertion pattern in which the insulator sequences are positioned close to the body enhancer. If the difference in *gypsy* orientation when inserted between the body enhancer and the promoter resulted in different *yellow* phenotypes in body tissues between the *mod(mdg4)*^{u1} mutants, the silencing effect in the absence of the Mod(mdg4)67.2 protein may only be local and therefore affect only the regulatory regions proximal to the Su(Hw) binding sites. To test this hypothesis, one transgenic line containing only the insulator sequences between wing enhancer and body enhancer of the *yellow* transgene was established as a homozygous *mod(mdg4)*^{u1} allele. As mentioned above, the *mod(mdg4)*^{u1} mutant line carrying a single *gypsy* retrovirus insertion between the wing enhancer and body enhancer showed black coloration in the body cuticle similar to wild-type transgenic lines (compare Figure 3.9C with Figure 3.9D). However, the *mod(mdg4)*^{u1} mutant line carrying only insulator sequences between the wing enhancer and body enhancer showed a variegated body phenotype (compare Figure 3.9G with Figure 3.9H). The difference in body

pigmentation between the two *mod(mdg4)*^{u1} mutant lines suggests that spreading of gene silencing originating from the insulator region was very limited and the local silencing could not reach the body enhancer in the transgenic line carrying the *gypsy* retrovirus due of the increased distance between the body enhancer and insulator region caused by the 7.5 kb *gypsy* insertion. However, the variegated bristles occurred in the *mod(mdg4)*^{u1} mutants of all transgenic lines, indicating that bristle enhancers are affected by silencing regardless of the difference in distance caused by the orientation of *gypsy* insertions(Table 3.1). This different response of enhancers to the silencing effect of the insulator in a *mod(mdg4)*^{u1} background remains unsolved, but results shown in next section suggest that insulators may exert some direct effect on transcription by directly interacting with promoters.

An upstream *gypsy* insulator increases transcriptional activity of enhancers.

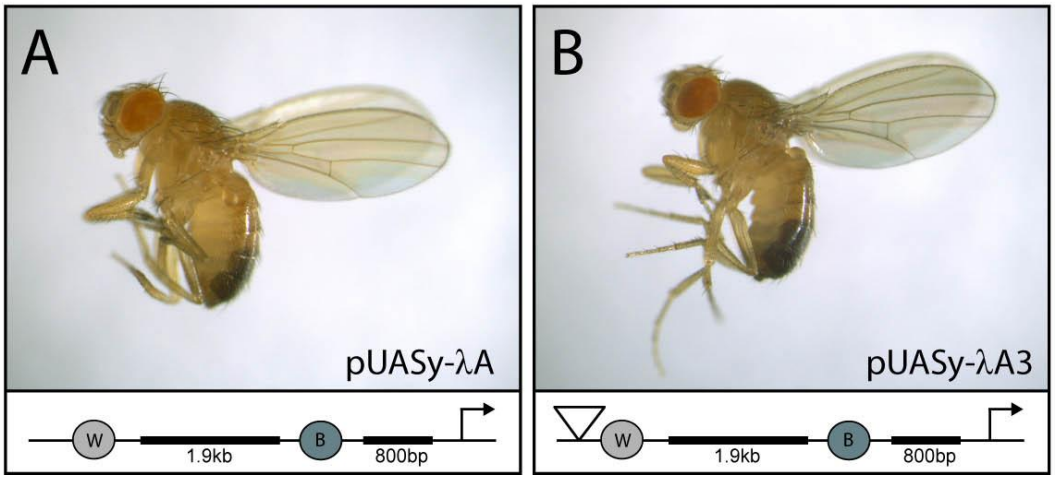
While the tandem repeat of two *gypsy* retroviruses with the same orientation inserted in the regulatory region of a *yellow* transgene blocked enhancer activity (Labrador et al., 2008), earlier studies showed that the tandem repeat of two *gypsy* insulators has no effect on enhancer-bypass (Kuhn et al., 2003; Majumder and Cai, 2003). Before considering that the discrepancy may result from other possible factors existing in the full-length *gypsy* retrovirus, we decided to test whether the increase in distance between the regulatory region and the promoter caused by the *gypsy* insertions affects the enhancer-promoter interaction in the *yellow* transgene. A new transformation vector containing *yellow* as a reporter gene has been developed, in

which the distance between wing and body tissue specific enhancers, as well as the distance between the enhancers and the *yellow* promoter were significantly increased by incorporating lambda DNA sequences to mimic the insertion of the longer *gypsy* retrovirus (Figure 2.1). The lambda DNA sequences used in these new plasmids included multiple cloning sites, allowing the directional cloning of DNA sequences between the wing and body enhancers and between the body enhancer and the promoter. In addition, FRT and loxP sequences flanking the multiple cloning sites also allow for the selective inversion of the cloned sequences once the transgene is inserted into a chromosome. Since this cloning vector was initially designed for analysis of insulator function, the Su(Hw) binding sites located downstream of the poly-adenylation site of the *yellow* gene were intentionally removed to prevent experimental interferences from endogenous insulators located within the transgene (Soshnev et al., 2008). This cloning vector was named pUAS*y*-λA. Transgenes obtained with the pUAS*y*-λA vector failed to induce expression levels that are normally obtained by *yellow* transgenes in the body cuticle and wing blades (Figure 3.10). The body enhancer, for example, was able to induce significant levels of expression, measured as coloration in adult cuticles, although these expression levels were lower than those observed in the wild-type gene or in other transgenes containing a wild-type copy of the *yellow* gene (Labrador and Corces, 2001; Labrador et al., 2008). The loss of pigmentation in wing blades was more severe, showing coloration practically identical to that found in wings from y^1 flies (Figure 3.10A). When a copy of the *gypsy* insulator was placed between the body enhancer and the promoter in this construct (pUAS*y*-λA1), the *gypsy* insulator

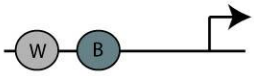
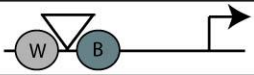
Figure 3.10 Transcriptional activation at the *yellow* promoter in pUAS*y*-λ transgenes by tissue specific enhancers is positively modulated by *gypsy* insulators located upstream of the enhancers.

Triangles denote *gypsy* insulator sites. Broken arrows correspond to the *yellow* gene promoter in the transgene. Thick lines correspond to lambda sequences. Wing blades and body cuticle tissue specific enhancers are shown as colored circles with their initials. Numbers in the wing and body phenotype columns correspond to the number of independent transgenic lines within a particular phenotypic category. Gray squares correspond to the average phenotype induced by the transgene. Examples of *yellow* phenotypic changes in pUAS*y*-λ transgenic flies in the absence (A) and presence (B) of an insulator upstream of the wing enhancer are shown at the bottom.

	Transgenes	No.	Phenotype									
			Wing					Body				
			1	2	3	4	5	1	2	3	4	5
pUAS y-λA		4	3	1							4	
pUAS y-λA 1		5	5						5			
pUAS y-λA 2		5	4	1							1	4
pUAS y-λA 3		20				2	18				1	19



completely blocked the activity of the body enhancer but no change in expression level was observed in wing blades. The same result was observed in the wing blades of transgenic lines that have an insulator between the wing enhancer and the body enhancer (pUAS*y*-λA2), indicating that wing enhancer activity could not reach the promoter in these constructs regardless of presence or absence of the insulator. However, the pUAS*y*-λA2 transgenic flies showed a significant enhancement of expression levels by the body enhancer, such that the body pigmentation was almost indistinguishable from that of wild-type flies. Since the level of pigmentation in the body cuticle was significantly increased once the insulator was placed upstream of the body enhancer, another construct placing the insulator upstream of both enhancers (pUAS*y*-λA3) was constructed to test whether an insulator upstream of the wing enhancer is also capable of rescuing the phenotypic defect in wing blades. Results show that pUAS*y*-λA3 transgenic flies carrying an upstream *gypsy* insulator have a phenotype indistinguishable from wild type flies in both body cuticle and wing blades(Figure 3.10B). To determine whether the lambda sequences introduced into the pUAS*y*-λA vector have a negative effect on the activity of the wing enhancer, transgenes using a plasmid vector identical to the pUAS*y*-λA but without the lambda DNA spacers were generated (pUAS*y*; Figure 2.1). In the pUAS*y* transgenic flies, wing and body enhancers showed almost normal activity (Figure 3.11A), indicating that the presence of lambda DNA spacers significantly reduced the activity of the wing enhancer in the pUAS*y*-λA transgenes. To test the effect that upstream insulators have on the activity of enhancers, an insulator was added upstream of the body enhancer in pUAS*y* transgene

	Transgenes	No.	Phenotype									
			Wing					Body				
			1	2	3	4	5	1	2	3	4	5
pUAS y		6				6					5	1
pUAS y-1		5		5								5

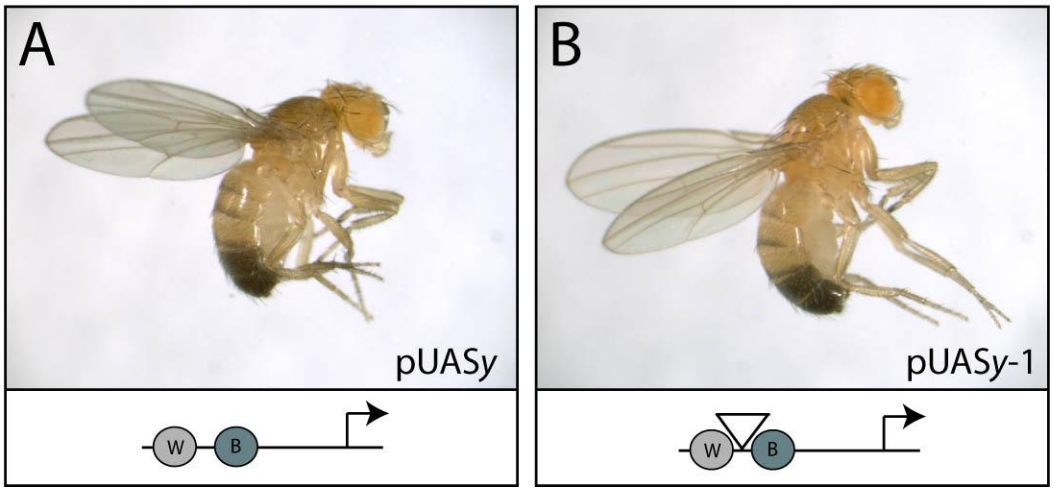


Figure 3.11 Wing and body enhancers activate transcription at the *yellow* promoter in pUAS*y* transgenes lacking lambda DNA.

Symbols are identical to Figure 3.10. Insulators block the activity of the wing enhancer in the pUAS*y*-1 transgene. Representative examples of flies containing pUAS*y* (A) and pUAS*y*-1 (B) transgenes are shown at the bottom.

(pUAS*y*-1). Analysis of different lines containing pUAS*y*-1 transgenes showed that the upstream insulator blocks the wing enhancer from activating transcription and clearly increases the levels of pigmentation in the body, in a manner similar to that of pUAS*y*-λA2 transgenes, when the insulator was placed upstream of the body enhancer (Figure 3.11B). These results show that the lambda spacers are in large part responsible for the lack of activity of wing enhancers in pUAS*y*-λA transgenes. The lack of expression in pUAS*y*-λA transgenes could be interpreted as if the lambda DNA is significantly reducing the activity of the wing enhancer by generating a repressive environment or by introducing distance between the wing enhancer and the promoter, which may diminish the ability of the wing enhancer to contact the promoter at the *yellow* gene.

The upstream *gypsy* insulator facilitates enhancer-bypass of paired insulators.

Since it has been determined that distal enhancers are able to bypass the enhancer-blocking activity of paired insulators, a model has been proposed in which the physical pairing of the insulators shortens the overall distance in the regulatory region upstream of the promoter and brings distal enhancers closer to promoters, facilitating eventual contacts between them (Cai and Shen, 2001; Muravyova et al., 2001). This model is supported by experimental results using the GAL4 binary system in which GAL4-UAS sequences activate transcription in the presence of GAL4 at a distal promoter only when two paired insulators are placed between the UAS and the promoter (Kyrchanova et al., 2008; Savitskaya et al., 2006). If distance caused by lambda spacers is a factor in the failure of the wing enhancer to activate transcription at

the *yellow* promoter in the pUAS*y*-λA transgenes, it was reasoned that two insulators placed between the wing enhancer and the promoter should facilitate transcriptional activation by interacting with each other and reducing the distance between the wing enhancer and the promoter. Transcriptional activation of the *yellow* promoter by the wing enhancer, however, was negative also for every single tested transgene that included paired insulators and lambda DNA sequences (pUASy-λA4 and pUAS*y*-λA5; Figure 3.12A and B). To further test the effect of distance in the ability of enhancers to bypass paired insulators, we generated an additional cloning vector in which the distance between the body enhancer and the promoter was further increased, while lambda sequences between the wing enhancer and the body enhancer were removed (pUASy-λB; Figure 2.1). Using the pUASy-λB vector, new transgenes (pUASy-λB1) were obtained by placing paired insulators between the body enhancer and the promoter (Figure 3.12). Under these new conditions, both the wing and body enhancers in pUASy-λB1 transgenes were still unable to bypass the enhancer-blocking activity of the insulators and failed to activate the *yellow* promoter. Since none of the transgenes tested so far in this study produced enhancers capable of bypassing paired insulators, we could not rule out the possibility that the combination of enhancers and insulators in our cloning vectors, independent of the lambda spacers, have not retained the capability of bypassing paired insulators. To determine whether enhancers in our transformation vectors were actually capable of bypassing the enhancer-blocking activity of paired insulators, we generated new transgenes with paired insulators and with no lambda sequences (pUASy2; Figure 3.12C). The pigmentation of wing blades in the pUAS*y*2

	Transgenes	No.	Phenotype									
			Wing					Body				
			1	2	3	4	5	1	2	3	4	5
pUAS y-λA 4	W 1.9kb B 800bp	7	7						7			
pUAS y-λA 5	W 1.9kb B 800bp	6	6									6
pUAS y-λB 1	W B 1.5kb	11	8	3					7	4		
pUAS y-2	W B	8		2	5	1			6	1	1	

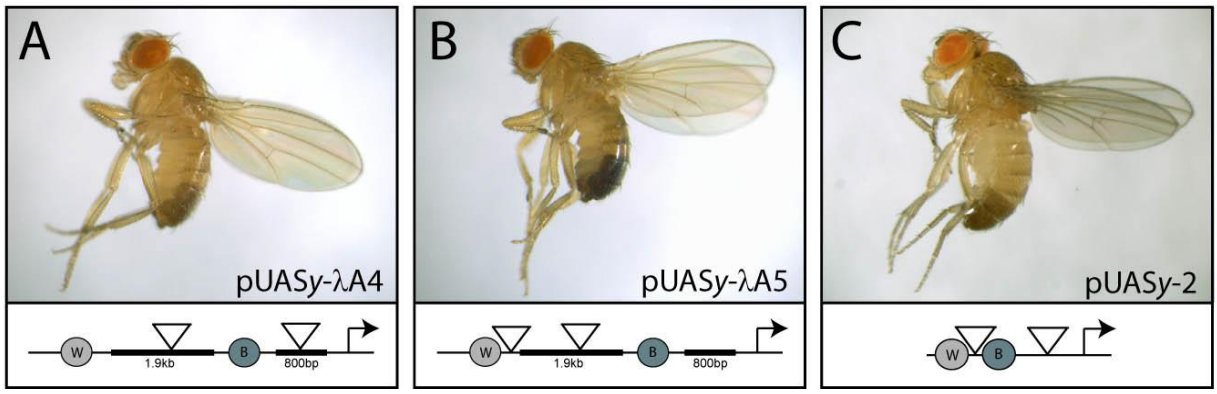


Figure 3.12 Enhancers fail to bypass the enhancer-blocking activity of paired insulators in pUAS*y*-λ transgenes.

Symbols are identical to Figure 3.10. The wing enhancer is only capable of significantly activating transcription by bypassing paired insulators in pUASy transgenes. In pUASy-λB1 transgenes a weak activation by distal enhancers in wing blades could be observed only in a small number of transgenic lines. Some representative examples of transgenic flies are shown at the bottom. The pUASy-λA4with lambda DNA (A) and pUASy-2 without lambda DNA (C) have the same insertion pattern of *gypsy* insulators, but show different wing phenotypes.

lines generally showed an intermediate coloration in wing pigmentation levels compared with that of the pUAS*y* and pUAS*y*1 lines, indicating that the wing enhancers systematically bypassed the enhancer-blocking activity of the paired insulators located downstream. Considering a model in which the interaction between insulators leads to the access of the promoter by the distal enhancer, the failure of enhancer-bypass in all these pUASy-λ lines containing two insulators suggests that the introduction of lambda DNA generates a repressive environment and does not support the notion that the distance introduced by the lambda spacers prevent enhancer activity from reaching the promoter.

The lack of insulator bypass by the wing enhancer in pUASy-λA4 was independent of insulator orientation. Because it had previously been demonstrated that distal enhancers only bypass the enhancer-blocking activity of *gypsy* insulators when paired *gypsy* proviruses are in opposite orientation (Figure 1.5B) (Labrador et al., 2008), some lines in pUASy-λA4 transgenic flies were induced to invert the orientation of one of the insulators by crossing the pUAS*y*-λA4 transgenic flies with flies expressing a transgenic copy of the Cre recombinase gene by using the site-specific recombination sites flanking the insulator sequences. The change in the relative orientation of two paired insulators did not affect the levels of *yellow* expression driven by the wing enhancer. Also, in the pUAS*y*2 lines, some lines contained two *gypsy* insulators in the opposite orientation and the other lines contained two *gypsy* insulators in the same orientation. However, all pUAS*y*2 lines with the different orientation of two insulators showed the bypass of wing enhancer activity, confirming that the orientation-dependent

enhancer-bypass does not result from the properties of the *gypsy* insulator but from the properties of the retrovirus.

Based on the previous observations showing that enhancers become stronger transcriptional activators after *gypsy* insulators are placed upstream of enhancers, we examined whether upstream *gypsy* insulators could also facilitate bypassing the enhancer-blocking activity of downstream paired insulators. New transgenes that are identical to pUAS*y*-λA4, pUAS*y*-λA5, and pUAS*y*-λB1, but contain a copy of the *gypsy* insulator upstream of the body or wing enhancer, were generated to this end (Figure 3.13). The pUAS*y*-λB2 and pUAS*y*-λB3 transgenes were derived from pUAS*y*-λB1, and the pUAS*y*-λA6 and pUAS*y*-λA7 transgenes were the derivatives of pUAS*y*-λA4 and pUAS*y*-λA5, respectively. While wing enhancers of pUAS*y*-λA4 and pUAS*y*-λA5 could not bypass two insulators and failed to activate *yellow* gene in wing tissue, pUAS*y*-λA6 and pUAS*y*-λA7, which have an insulator upstream of wing enhancer, successfully activated the *yellow* transgene to levels similar to those of the wild-type *yellow* gene in the wing tissue (compare Figure 3.12A and B with Figure 3.13A and B). Also, while both the wing and body enhancer of the pUAS*y*-λB1 failed to induce *yellow* gene expression, the addition of an insulator upstream of the body enhancer helped the body enhancer to activate the *yellow* transgene (pUAS*y*-λB2) and the addition of an insulator upstream of the wing enhancer caused expression of the *yellow* transgene in wing and body tissues (pUAS*y*-λB3). These results show that upstream insulators also help enhancers to bypass the enhancer-blocking activity of paired insulators.

One possibility, given that these transgenes have three insulators alternating with

	Transgenes	No.	Phenotype									
			Wing					Body				
			1	2	3	4	5	1	2	3	4	5
pUAS y-λB 2	W B 1.5kb	3	3									3
pUAS y-λB 3	W B 1.5kb	4			1	1	2		1	1	2	
pUAS y-λA 6	W 1.9kb B 800bp	4					4		4			
pUAS y-λA 7	W 1.9kb B 800bp	6		1	2	2	1					6
pUAS y-λA 8	W 1.9kb B 800bp	6	6						6			

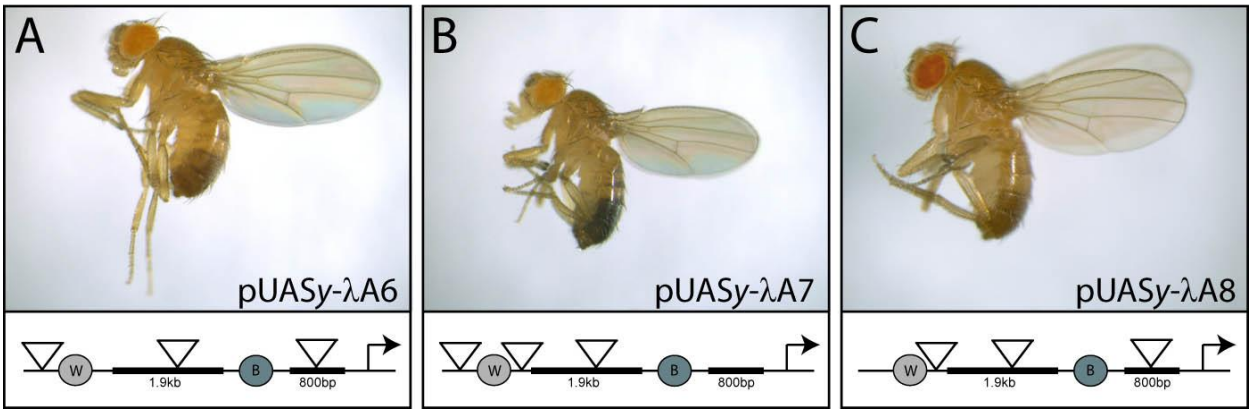


Figure 3.13 An upstream *gypsy* insulator strongly facilitates bypassing of downstream paired insulators by the body and the wing enhancers.

The level of pigmentation in wing and body tissue were evaluated in transgenic lines in which one *gypsy* insulator is placed upstream of tissue specific enhancers of *yellow* transgenes described in Figure 3.12. Some representative examples of transgenic flies are shown at the bottom. Compared with the pUASy-λA4 (Figure 3.12A) and the pUASy-λA5 (Figure 3.12B), the significant increase of *yellow* gene expression in wing tissue of pUASy-λA6 (A) and pUASy-λA7 (B) could be observed, respectively.

enhancers, is that bypass of enhancer blocking is the result of the interaction between the three insulators, and is actually independent of the location of the third insulator relative to the enhancer. To test this idea, a third insulator downstream of the wing enhancer in the pUAS*y*-λA4 transgene was placed to obtain a new transgene named pUAS*y*-λA8 (Figure 3.13C). The pUAS*y*-λA8 transgenic lines showed that the wing enhancer was unable to activate transcription, ruling out the possibility that only interactions between insulators are sufficient to facilitate transcriptional activation. This result supports the notion that activation by enhancer bypass also requires a third insulator upstream of the enhancer in these transgenes.

Classical position effects alone cannot explain the modulation of transcriptional activity of enhancers by the *gypsy* insulator.

Together, these results suggest that insulators function as transcriptional modulators that affect downstream enhancers in a positive manner. However, we could not exclude the possibility that the level of transcriptional activation observed in transgenes with insulators upstream of the enhancers is due to the classical boundary function of insulators, which protects transgenes against position effects. Hence, the phenotypic distribution of wing blades and body cuticle in all transgenes obtained in this study using pUAS*y*-λ derived vectors was compared (Tables 3.2). None of the 28 independent transgenic lines lacking an upstream insulator show significant levels of wing pigmentation. However, when a *gypsy* insulator is placed upstream of the wing enhancer, 30 out of 34 lines analyzed showed close to normal levels of wing

Table 3.2 Summary of *yellow* phenotypes in wing and body tissues of the pUASy-λ transgenic lines.

A Comparison of *yellow* expression by the wing enhancer in transgenic lines with or without an insulator upstream of the enhancer.

Upstream insulator	Downstream insulator	*pUASYλ* transgenic lines	Wing 1	2	3	4	5
Without	No	pUAS*y*-λA	3	1			
	Two	pUAS*y*-λA4	7				
		pUAS*y*-λA5	6				
		pUAS*y*-λB1	8	3			
		Total	24	4	0	0	0
With	No	pUAS*y*-λA3				2	18
	Two	pUAS*y*-λB3			1	1	2
		pUAS*y*-λA6					4
		pUASy-λA7		1	2	2	1
		Total	0	1	3	5	25

B Comparison of *yellow* expression by the body enhancer in transgenic lines with or without an insulator upstream of the enhancer.

Upstream insulator	Downstream insulator	*pUASYλ* transgenic lines	Body 1	2	3	4	5
Without	No	pUAS*y*-λA				4	
	Two	pUAS*y*-λB1		7	4		
		Total	0	7	4	4	0
With	No	pUAS*y*-λA2				1	4
		pUAS*y*-λA3					10
		pUAS*y*-λA5					6
		pUAS*y*-λA7					6
	Two	pUAS*y*-λB2					3
		pUAS*y*-λB3		1	1	2	
		Total	0	1	1	3	29

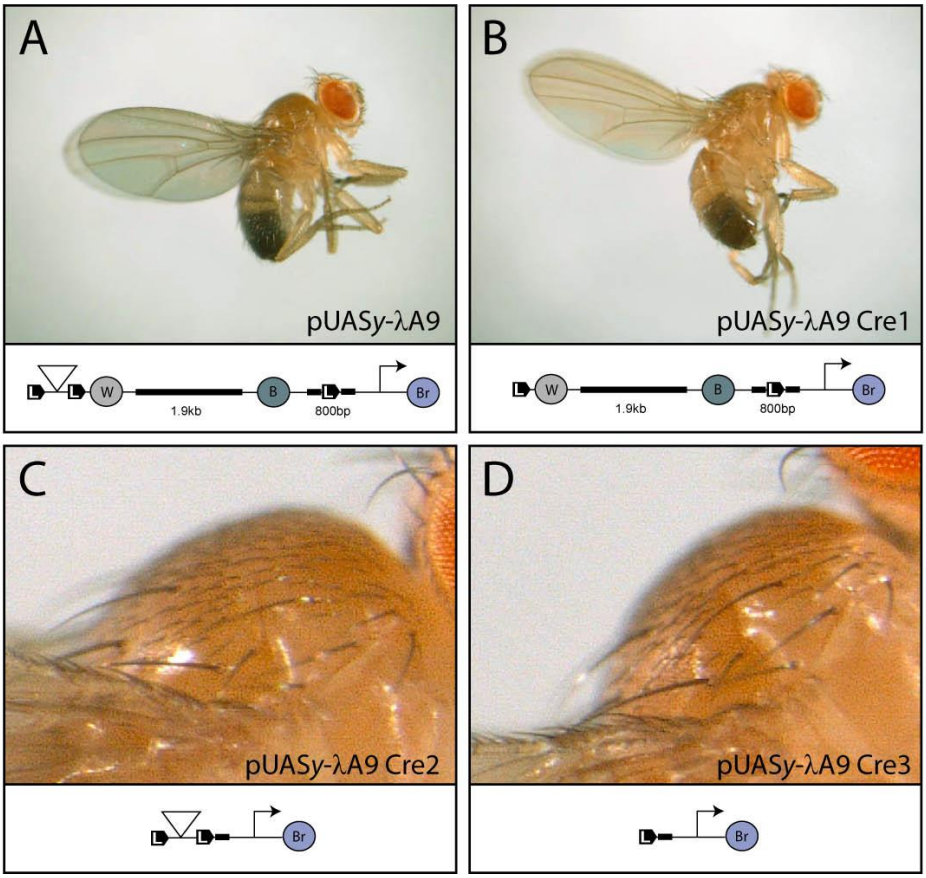
pigmentation (Table 3.2A). A pattern similar to the distribution of pigmentation in wing blades is also shown in body cuticles (Table 3.2B), although the difference in pigmentation distribution according to the presence and absence of the upstream insulator is not as dramatic as that observed in wing tissue, probably because of the relative difference in distance between the promoter and the two tissue-specific enhancers. Considering that classical position effects are normally associated with a gradation in the level of expression of the transgene according to the insertion site and are manifested in a positive and a negative manner as well, the persistent lack of wing pigmentation in all pUAS*y*-λ transgenic lines devoid of insulator sequences upstream of the wing enhancer and the phenotypic homogeneity among lines supports that the inability of the wing enhancer to activate *yellow* in the pUAS*y*-λ transgenes is not due to traditional repressive position effects.

To further investigate whether negative position effects were responsible for the lack of expression of the pUAS*y*-λ transgenes, a new plasmid (pUAS*y*-λA9) containing a tandem repeat of loxP sites at both ends of an insulator, located upstream of the wing enhancer, was constructed (Figure 3.14). Crossing the pUAS*y*-λA9 transgenic flies with Cre-expressing lines caused three different types of transgenic flies (pUAS*y*-λA9 Cre1, pUAS*y*-λA9 Cre2 and pUAS*y*-λA9 Cre3). In the pUAS*y*-λA9 Cre1 lines, only the upstream insulator was removed by site-specific recombination of loxP sites at both ends of the upstream insulator. The pUASy-λA9 Cre2 and pUASy-λA9 Cre3 lines were obtained by recombination between the loxP sites found at both ends of the insulator located upstream of the wing enhancer and the loxP sites between the body enhancer

Figure 3.14 Removing the upstream *gypsy* insulator by site-specific recombination eliminates the ability of the wing enhancer to activate transcription at the *yellow* promoter.

The target sites of Cre recombinase for site-specific recombination in the pUASy-λA9 transgene are marked as closed pentagon boxes with the initial of LoxP (L). Circle labeled Br represents bristle enhancer and other designations are as in Figure 3.10. **(A - D)** While removal of the *gypsy* insulator upstream of the wing enhancer by site-specific recombination causes significant decrease in the expression of the *yellow* gene in wing tissue (Compare A with B), *yellow* phenotypes in bristles of the pUASy-λA9 Cre 2 and Cre 3 transgenic lines do not show any difference regardless of the presence or absence of the *gypsy* insulator (Compare C and D).

	Transgenes	No.	Phenotype									
			Wing					Body				
			1	2	3	4	5	1	2	3	4	5
pUAS y-λA9	W 1.9kb B 800bp Br	8				1	7				1	7
pUAS y-λA9 Cre1	W 1.9kb B 800bp Br	6	6							1	5	
pUAS y-λA9 Cre2	Br	3	3					3				
pUAS y-λA9 Cre3	Br	6	6					6				



and the promoter. In these two lines, the upstream regulatory region of the pUASy-λA9 transgenes, including the lambda sequences, was excised. While the *yellow* promoter and bristle enhancer were adjacent to the genomic environment in the pUASy-λA9 Cre3 lines, an insulator was flanking the transgene in the pUASy-λA9 Cre2 lines. Elimination of the upstream insulator from 6 independent transgenic lines (pUAS*y*-λA9 Cre1) led to a decrease in their pigmentation in body and wing tissues to a level similar to that of pUASy-λA transgenes, which originally lacked the insulator. This phenotypic reversion seems to support that insulators upstream of the enhancers play a role as boundaries to protect transgenes against a negative environment. However, if this is true, the pUASy-λA9 Cre3 lines should show yellow or variegated bristle phenotypes because the *yellow* promoter and bristle enhancer are exposed to the negative environment. The degree of bristle pigmentation and variegation in the thorax and head has been used before as a marker for position effects (Golovnin et al., 2005; Kurshakova et al., 2007), indicating that the bristle enhancer located downstream of promoter is sensitive to genomic context. Nevertheless, like the pUASy-λA9 Cre2 lines in which the insulator protects the *yellow* promoter and bristle enhancer against the genomic environment, the pUASy-λA9 Cre 3 lines have a wild-type level of pigmentation in the bristles with no variegation, suggesting that negative position effects do not have a major role in the phenotypic determination of the pUASy-λ transgenes used in this study. In addition, when a wide range of eye color pigmentation was compared among all transgenes, the position effects influencing the *white* transgene expression were not correlated with the phenotype induced by the *yellow* transgene.

The minimal *OMB* wing enhancer cannot bypass paired *gypsy* insulators

We showed above that a *gypsy* insulator placed upstream of an enhancer increases transcriptional activity of the enhancer facilitating enhancer bypass of the enhancer-blocking activity of paired insulators downstream of the enhancer. A question that arises here is whether the insulator upstream of the enhancer interacts with insulator(s) downstream of the enhancer. For example, the pUAS*y*-λA6 transgenic lines showed *yellow* gene expression in wing blades. On the basis of the model which has been suggested by previous studies (Figure 1.3C), it can be assumed that two insulators downstream of the wing enhancer interact with each other and the wing enhancer bypasses the paired downstream insulators in pUAS*y*-λA6. However, if the upstream insulator interacts with one of the two downstream insulators, a new independent chromatin loop domain forms, leaving the wing enhancer in the loop domain and the *yellow* promoter outside. In this chromatin structure, the wing enhancer cannot activate the *yellow* promoter. Hence, any selective interaction between insulators may be influenced by whether they are placed upstream or downstream of the enhancer, although we showed through the GAL4-UAS system that there is fundamentally no preferential binary interaction between three insulators. We inferred that enhancers, which are activated in specific tissues and at particular times of development, interrupt the interaction between insulators placed upstream and downstream of the enhancers. To test this hypothesis, the transformation vector p*CI-OMB* RG was constructed. The plasmid construct contained EGFP and DsRed reporter

genes, two *gypsy* insulators, and the *optomotor-blind* [(Secombe and Parkhurst) wing enhancer and the *cubitus interruptus* (*ci*) enhancer, two enhancers which induce specific expression patterns in wing imaginal discs. *optomotor-blind* [(Secombe and Parkhurst), which encodes a T-box domain transcription factor (Pflugfelder et al., 1992) as a downstream target of Dpp signaling (Grimm and Pflugfelder, 1996), is expressed regionally in anterior and posterior wing disc cells along the anterior/posterior (A/P) axis, and this expression pattern is under the control of the *omb* w enhancer (Sivasankaran et al., 2000). *cubitus interruptus* (*ci*), encoding a transcription factor involved in Hedgehog signaling (Alexandre et al., 1996; Hepker et al., 1997), is expressed throughout the anterior compartment in wing discs (Eaton and Kornberg, 1990). Therefore, the expression of *ci* and *omb* partially overlap in an area of the anterior compartment, indicating that two enhancers are simultaneously activated in disc cells in that area. In the p*CI-OMB* RG construct, the *ci* enhancer and the EGFP reporter gene were separated by the intervening lambda DNA spacer. One insulator sequence was placed downstream of the *ci* enhancer and the other insulator was upstream of EGFP promoter. The *omb* wing enhancer and DsRed reporter gene were then inserted in the lambda DNA between the *ci* enhancer and EGFP reporter gene (Figure 3.15A). It was expected that the interaction between the two insulators would lead to bypass of the *ci* enhancer to activate EGFP expression in the anterior compartment and the *omb* enhancer to activate only DsRed expression in the chromatin loop domain formed by the insulator interaction (Figure 3.15B). If the active *omb* enhancer does not interrupt the interaction between insulators downstream and upstream of the *omb* enhancer,

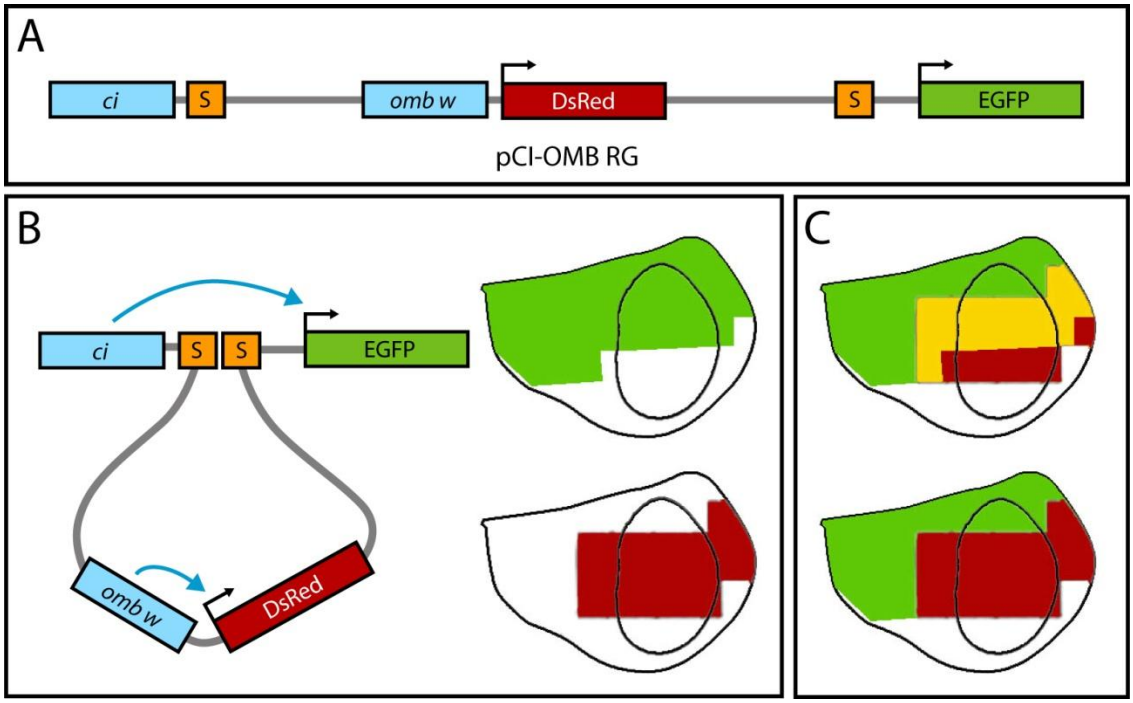


Figure 3.15 Schematic drawing of p*CI-OMB* RG transgene.

(A) Two reporter genes, EGFP and DsRed, are illustrated as green and red boxes, respectively, and broken arrows mark promoter regions. The blue boxes indicate *ci (cubitus interruptus)* and *omb* (*optomotor-blind*) wing enhancers, and *gypsy* insulator sequences are depicted as orange boxes with S. **(B)** The expected expression patterns of EGFP and DsRed by the *ci* enhancer and *omb* wing enhancer in wing imagnial discs, respectively, when two insulators interact. Blue arrows indicate reporter gene activation by enhancers. **(C)** Overlapped region of expression of both reporter genes by *ci* and *omb* wing enhancers is shown in yellow (top). Expected expression pattern when active *omb* wing enhancer interrupts the interaction between the upstream and downstream insulators in the overlapped region (bottom).

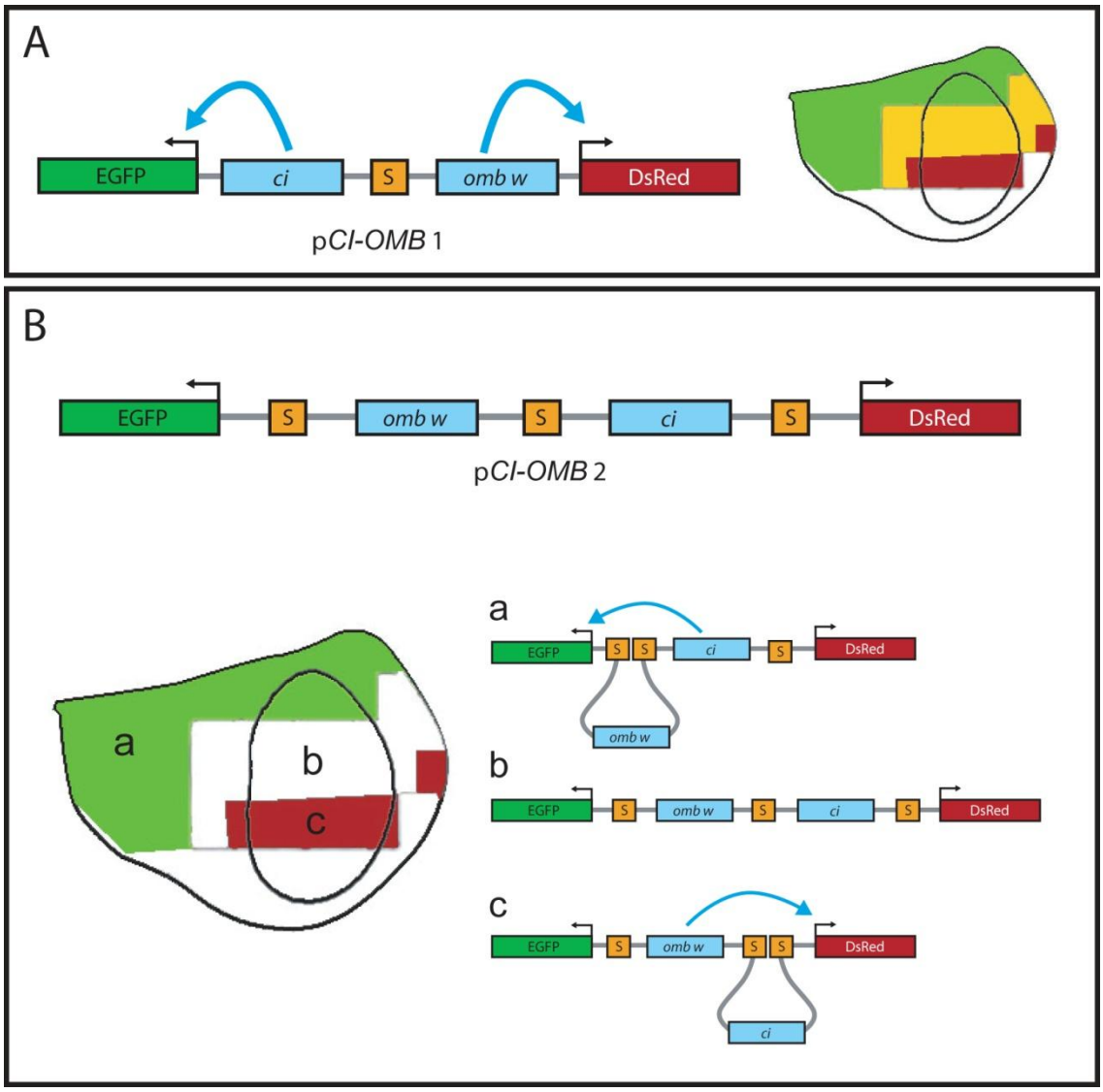
EGFP expression in the area of the anterior compartment in which DsRed is expressed would show co-localization of the two reporter gene expression signals which would appear as a yellow color in the merged image. If the active *omb* enhancer does interrupt the interaction, EGFP expression in the area where DsRed is expressed would not be observed because the *ci* enhancer cannot bypass the two insulators without insulator interactions (Figure 3.15C). Unexpectedly, EGFP expression by the *ci* enhancer was not observed in any anterior compartment cells in the wing disc, whereas DsRed was expressed specifically in the local area along the A/P axis by the *omb* enhancer.

The *ci* gene is located on chromosome 4, the smallest of the autosomes, in *D. melanogaster.* Chromosome 4 has an unusual chromatin organization exhibiting not only several characteristics typical of heterochromatin, but also a gene density typical of euchromatin (Riddle and Elgin, 2006). Little is known about the regulation mechanism of genes in the particular heterochromatin environment of chromosome 4, although it is known that certain genes located in the heterochromatin region require a heterochromatin environment for their expression (Eberl et al., 1993; Weiler and Wakimoto, 1998). However, transgenic lines containing the *ci* enhancer and *lacZ* reporter gene showed normal expression patterns (Schwartz et al., 1995). Therefore, the lack of expression in wing disc does not seem to be due to the inability of the *ci* enhancer to function in regions of euchromatin. One possibility is that the failure of EGFP expression by the *ci* enhancer in p*CI-OMB* RG may be due to the lambda DNA spacer added between the *ci* enhancer and EGFP reporter gene, similar to the failure of *yellow* gene activation by the wing enhancer in pUAS*y*-λ transgenic lines. Two new

constructs (p*CI-OMB* 1 and p*CI-OMB* 2) were prepared to test this possibility. The p*CI-OMB* 1 transgene was constructed to determine whether the *ci* enhancer was actually capable of inducing the expression of the EGFP reporter gene when the *ci* enhancer is placed next to the promoter of the EGFP reporter gene. In this construct, EGFP and DsRed reporter genes are placed next to the *ci* enhancer and the *omb* wing enhancer, respectively. Also, a single *gypsy* insulator, which would act as a boundary to prevent improper activation of reporter genes by the enhancer on the other side, was placed between the *ci* enhancer and the *omb* wing enhancer (Figure 3.16 A). The p*CI-OMB* 2 construct was prepared to utilize the ability of the insulator to stimulate enhancer activity when it is placed upstream of the enhancer. In the *pCI-OMB* 2 transgene, the EGFP and DsRed reporter genes were placed in opposite orientation and three insulators and lambda DNA were inserted between the two reporter genes. The *ci* enhancer was inserted between the first insulator, distal to the EGFP reporter gene, and the middle (2^{nd}) insulator. The *omb* wing enhancer was placed between the third insulator, distal to DsRed, and a middle insulator. Based on the *yellow* gene activation by the wing enhancer observed in the pUAS*y*-λA6 transgenic lines (Figure 3.16B), this arrangement should result in the activation of EGFP by the *ci* enhancer which bypasses the interaction of the 2^{nd} and 3^{rd} insulators, and the activation of DsRed by the *omb* wing enhancer which bypasses the interaction of the 1^{st} and 2^{nd} insulators. In both p*CI-OMB* 1 and p*CI-OMB* 2 lines, expression of EGFP by the *ci* enhancer was not induced. The failure of EGFP expression by the *ci* enhancer in p*CI-OMB* 1, where the *ci* enhancer is placed close to the EGFP reporter gene, indicates that the enhancer could not carry out

Figure 3.16 Schematic drawing of p*CI-OMB* 1 and 2 transgenes and expected expression patterns of reporter genes in wing imaginal disc by insulator interaction.

Symbols are identical to Figure 3.14 **(A)** In p*CI-OMB* 1, enhancer-blocking activity of the *gypsy* insulator between two enhancers will lead to EGFP expression by *ci* enhancer and the DsRed expression by *omb* wing enhancer (left). Expected expression patterns in wing imaginal discs [right]. **(B)** Scheme of p*CI-OMB* 2 transgene construct (top). In region a of wing imagnial discs where EGFP is expressed, active *ci* enhancer disrupts the interaction between the upstream and downstream insulators, and so only two insulators upstream of the *ci* enhancer interact. In this configuration, the active *ci* enhancer will activate the EGFP reporter gene. In region b where EGFP and DsRed reporter genes are not expressed, both active enhancers disrupt the interaction among three insulators, and so the two enhancers will not be able to activate EGFP and DsRed reporter genes. In region c where DsRed is expressed, the active *omb* wing enhancer allows only two downstream insulators to interact. In this configuration, the active *omb* wing enhancer will activate the DsRed reporter gene.


A
EGFP
ci
S
omb w
DsRed
pCI-OMB 1
B
EGFP
S
omb w
S
ci
S
DsRed
pCI-OMB 2
a
b
c
a
EGFP
S
S
ci
S
DsRed
omb w
b
EGFP
S
omb w
S
ci
S
DsRed
c
EGFP
S
omb w
S
S
DsRed
ci

normal function. In addition, while expression of DsRed by the *omb* wing enhancer was induced in p*CI-OMB* 1 lines, it was not in p*CI-OMB* 2 lines. This indicates that the *omb* wing enhancer cannot bypass the two upstream insulators placed between DsRed and the enhancer. The *omb* wing enhancer that was used for our transgene constructs was the minimal size able to lead to a normal expression pattern when the enhancer regionis placed next to the reporter gene. It is possible that the *omb* wing enhancer, located approximately 27kb upstream of the *omb* transcriptional unit in the genome, may require additional sequences in the intervening region between the enhancer and the *omb* promoter for long-range interaction with the promoter and to bypass paired insulators.

CHAPTER IV

DISCUSSION

Orientation-dependent neutralization of insulator function requires factors other than the *gypsy* insulators per se.

Ever since the existence of insulators was first described, it has been suggested that insulators may play an important role in the compartmentalization of the genome into functional chromatin domains (Bell et al., 1999; Kellum and Schedl, 1992). The discovery of insulator enhancer-blocking and heterochromatin-boundary activities strengthened a notion that insulators may establish physical limits on the activity of regulatory factors on the chromatin fiber. The findings that *gypsy* insulators as well as scs and scs' in *Drosophila* could physically interact, elicited a mechanism by which the intervening DNA could represent an independent chromatin domain, insulated from outside regulatory factors (Blanton et al., 2003; Cai and Shen, 2001; Mongelard and Corces, 2001; Muravyova et al., 2001). These findings were somehow obscured by the realization that enhancers as well as repressors could bypass the insulator activity of two or more interacting insulators (Cai and Shen, 2001; Comet et al., 2006; Kuhn et al., 2003; Muravyova et al., 2001). Since enhancers or repressors could bypass the boundaries created by the insulators themselves, this observations left open the question of how insulators could actually define the boundaries between domains if their physical interactions lead enhancers or repressors to overcome their activity. As in

previous reports (Muravyova et al., 2001), our results show that tissue-specific enhancers can activate the distal *yellow* promoter when three *gypsy* retroviruses are interposed between the tissue-specific enhancers and the *yellow* promoter. However, the relative orientation between three *gypsy* retrovirues appears to be the factor that determines whether the enhancers can activate the promoter or is blocked by the insulators. The orientation-dependent phenotypic changes were also observed when two *gypsy* retroviruses were inserted between tissue-specific enhancers and *yellow* promoter (Labrador et al., 2008). Considering that the relative orientation of two or more *gypsy* retroviruses inserted between tissue-specific enhancers and *yellow* promoter play a crucial role for enhancer-bypass while the relative orientation of the *gypsy* insulator has no effect, other factors that affect enhancer-bypass, in addition to the insulator, may exist in the *gypsy* retrovirus.

Compared with 350 bp *gypsy* insulator, a possible factor is the extended length of the *gypsy* retrovirus spanning approximately 7.5 kb that might result in the increased distance between enhancer and promoter, according to the relative orientation of *gypsy* insertions and their interactions. If this is the case, increasing the distance between tissue-specific enhancers and *yellow* promoter should lead to a reduced degree or an attenuation of *yellow* gene activation. However, previous studies have shown that the phenotype of the y^2 mutant allele, caused by the insertion of a single *gypsy* retrovirus between the tissue-specific enhancers and the promoter of the endogenous *yellow* gene, is rescued to a wild-type phenotype in a *su(Hw)* mutant background (Gerasimova et al., 1995; Ramos et al., 2006), indicating that the *yellow* gene tissue-specific

enhancers can activate the *yellow* promoter even if the normal distance to the promoter is increased by 7.5 kb. Interestingly, we have observed that the pigmentation levels of the wing blades and body cuticle in a *su(Hw)* mutant background decreases with the number of *gypsy* insertions between tissue-specific enhancers and the promoter of a *yellow* transgene (Figure 3.7 and 3.8). Nevertheless, this decrease in *yellow* gene expression level seems not to be due to a mere increase of distance between the enhancers and the promoter. Although there is no difference in distance between the body enhancer and promoter in both two and three *gypsy* insertion transgenic lines, the body cuticles of *su(Hw)*$^{-}$ flies carrying three *gypsy* insertions was more significantly impaired in pigmentation level than body cuticles of *su(Hw)*$^{-}$ flies carrying two *gypsy* insertions (compare B and D with F and H in Figure 3.8). Therefore, phenotypic changes dependent on the relative orientation of two or more *gypsy* retrovirus insertions may also be due to other factors emanating from the *gypsy* retrovirus itself rather than by the differences in distance between enhancer and promoter caused by interactions between insulators.

A possible interpretation of the orientation-dependent effects of multiple *gypsy* insertions is that specific sequences may exist within the *gypsy* retrovirus, which are independent from the *gypsy* insulator sequence and that can negatively regulate enhancer-bypass and may operate in an orientation-dependent manner. There are some examples that specific properties of some insulators are a combination of distinct elements and therefore can be separated. Similar to the *gypsy* retrovirus, the *Idefix* retrotransposon also functions as an insulator, and two *Idefix* insertions result in the loss

of insulator activity (Conte et al., 2002a). However, while the insulator region of *gypsy* retrovirus is located downstream of the 5' LTR and pairing of two *gypsy* insulator regions is enough for enhancer bypass, the insulator of *Idefix* is localized within a 470-bp fragment corresponding to the U3 region of the LTR and requires the 5' UTR sequences downstream of the LTR for neutralization of *Idefix* insulator function (Brasset et al., 2007). 5' HS4 of the chicken β-globin locus and Locus Control Region (LCR) of the mouse TCR-α locus have boundary properties as well as enhancer blocking properties, but the boundary elements are independent of CTCF and can be separated from the enhancer blocking element (Gomos-Klein et al., 2007; Pikaart et al., 1998; Prioleau et al., 1999)

It is tempting to speculate that the 5' LTR region of *gypsy* retrovirus may include sequences that can modulate enhancer-bypass. Several arguments can be used that point to this possibility: First, distinct DNA elements that lead to the composite properties of insulators mentioned above are generally localized close to each other, and the 12 Su(Hw) binding sites as *gypsy* insulator are located next to the 5' LTR of *gypsy* retrovirus. Second, the LTR regions of some retrotransposons have been reported to affect neighboring gene expression through promoter competition between the promoter in 5' LTR region and neighboring gene promoters (Coney and Roeder, 1988; Conte et al., 2002b). In the case of the *Idefix* retrotransposon, promoter competition especially depends on the orientation of *Idefix* promoter, probably due to selective enhancer blocking according to the relative arrangement of promoter and insulator in the *Idefix* LTR region (Conte et al., 2002b). Third, LTRs are one of the main

targets for short RNAs-mediated silencing on the eukaryotic genome (de Wit et al., 2005; Martens et al., 2005; Mikkelsen et al., 2007). Moreover, the 5' UTR region in the 5' LTR of *gypsy* retrovirus was precisely characterized to be targeted by the RNA-mediated Flamenco repression (Sarot et al., 2004). Therefore, as a number of factors related to gene silencing or heterochromatin formation are placed close to *gypsy* insulators, enhancer bypass may be modulated by the relative orientation between insulator and these silencing factors.

Orientation-dependent insulator neutralization may also be achieved by endogenous insulators which are broadly distributed through the genome, via co-localization of two or more different endogenous insulators, as it occurs when two or more *gypsy* retrovirus are integrated in close proximity. Kyrchanova *et al* showed that when two copies of a composite DNA fragment, in which four consecutive binding sites for dCTCF protein and four consecutive binding sites for Su(Hw) protein are next to each other, were placed between a regulatory element and a distal promoter, bypass of the regulatory element was dependent on the relative orientation of the two composite DNA fragments (Kyrchanova et al., 2008). From this result, the authors suggested that the binding of at least two different insulator proteins might be essential for functional interactions between insulators in a manner dependent on the orientation. The genome-wide mapping of binding sites for *Drosophila* insulator proteins has revealed that several insulator proteins are frequently co-localized (Bushey et al., 2009; Negre et al., 2010). Accordingly, the distribution of binding sites for diverse insulator proteins as well as frequent co-localization of these binding sites in large stretches of chromosomal DNA

may not be random but may have an active role in transcription regulation through functional mechanisms that lead to selective interactions between regulatory elements and promoters according to relative arrangement of co-localized binding sites of insulator proteins.

Activated enhancers may disrupt the interaction between insulators flanking the enhancer.

Kuhn *et al* reported that tissue-specific enhancers of the *yellow* gene could not bypass three *gypsy* insulators placed in succession between the tissue-specific enhancers and the *yellow* promoter, suggesting that *gypsy* insulator interactions may occur only as pairs (Kuhn et al., 2003). On the other hand, Savitskaya et al showed that three *gypsy* insulators in the *yellow* reporter gene could not block enhancer activities (Savitskaya et al., 2006). This result was supported by genetic analysis using the model system of the *mini-white* gene, and the discrepancy from the earlier study was suggested to be due to the difference of the topological limit that may affect interaction of three insulators depending of the distance between insulators (Kostyuchenko et al., 2008). Nevertheless, these results do not explain whether the enhancer bypass results from the interactions among all three insulators or the preferential interaction of only two insulators. Consistent with these studies suggesting that neutralization of the enhancer-blocking property of insulators may take place after three insulators are placed between an enhancer and a promoter, we observed that wing enhancers activate the *yellow* promoters in class II and class III of our transgenic lines containing three *gypsy*

retrovirus insertions (Figure 3.2B and C). Moreover, by using the UAS-GAL4 system, we provided direct evidence that interactions among three *gypsy* insulators can lead to enhancer-bypass. While crossing of P[G4(S)E(S)D] and P[G4(S)(S)ED] transgenic lines with *GMR*-Gal4 lines resulted in strong expression of only one of the two reporter genes, according to the position of the insulator placed upstream of the reporter genes (Figure 3.4B and C), P[G4(S)(S)E(S)D] showed similar expression levels of both EGFP and DsRed reporter genes (Figure 3.4D). Although we did not investigate whether EGFP and DsRed reporter genes in P[G4(S)(S)E(S)D] lines are simultaneously expressed in a single cell, the specific expression pattern of P[G4(S)(S)E(S)D] lines is distinct from other transgenic lines containing two insulators in eye tissue and indicates that all three insulators actively participate in their interaction.

How, then, do multiple *gypsy* insulators interact? The *gypsy* insulator consists of 12 repeated binding sites of the Su(Hw) protein (Spana et al., 1988), therefore, our results suggest that multiple *gypsy* insulators may be able to interact with themselves until the interactions encounter a spatial limit. Unlike the *gypsy* insulator, endogenous insulators found in the *Drosophila* genome do not have a tightly clustered number of Su(Hw) binding site repeat with most endogenous insulators having a single Su(Hw) consensus binding site (Adryan et al., 2007). Nevertheless, the formation of insulator bodies consisting of only 25-30 dots per diploid interphasic cell nucleus shows that interactions among multiple endogenous insulators may occur in the *Drosophila* genome (Gerasimova et al., 2000; Gerasimova and Corces, 1998; Pai et al., 2004). The Mod(mdg4) protein belongs to the BTB-ZF family (Buchner et al., 2000), and some

proteins of the BTB-ZF family, including the GAGA factor, have been reported to be able to perform self-oligomerization as well as dimerization (Dong et al., 1996; Espinas et al., 1999). In addition, CP190 also contains a BTB/POZ domain (Pai et al., 2004), and the BTB domains of both Mod(mdg4) 67.2 and CP190 may mediate multiple endogenous *gypsy* insulator interactions.

Contrary to the suggested functional role of insulators to regulate improper activation of non-cognate genes by promiscuous enhancer activity in the genome, the enhancer may control the unlimited interactions of multiple insulators, which could result in improper gene activation or repression. We observed that while class I and II transgenic flies have yellow coloration in body cuticles (Figure 3.2 A and B), class III transgenic flies show black pigmentation in the body (Figure 3.2C). If three insulators in class I, II, and III transgenic flies are interacting, leading to the formation of two chromatin loops, body enhancers and *yellow* promoters are placed inside and outside of distinct loop domains in all transgenic lines, respectively. Therefore, the activation of the *yellow* gene by the body enhancer in the class III flies conflicts with a hypothesis that the formation of chromatin loop domains by interaction of insulators blocks the interaction between enhancer and promoter separately placed inside and outside the loop domain (Gerasimova and Corces, 1996; Kuhn and Geyer, 2003; Valenzuela and Kamakaka, 2006; Wallace et al., 2009). These results observed in transgenic lines with three *gypsy* retroviruses are also consistently observed in all pUASy-λ transgenic lines with three *gypsy* insulators. Similar to the case of the body enhancer in class III (Figure 3.2C), enhancers placed between one upstream insulator and two downstream

insulators activate the distal downstream promoter. For example, this occurs with body enhancers of pUASy-λB2, wing and body enhancers of pUASy-λB3, and wing enhancers of pUASy-λA6 and λA7 (Figure 3.13). On the contrary, like the body enhancer in class I and II (Figure 3.2A and B), enhancers that have a single downstream insulator and two upstream insulators, such as the body enhancers of pUASy-λA6 and λA8 (Figure 3.13), fail to activate the promoter. A persuasive explanation is that an activated enhancer may interrupt the interaction between the upstream and downstream insulators. For example, in wing tissues of class II and class III, the wing enhancer is activated, whereas the body enhancer is inactive. Therefore, the interaction of three insulators between the wing enhancer and promoter lead to wing enhancer bypass and expression of the *yellow* gene in wing tissues of both class II and III . On the other hand, in body tissues of class II and III, the wing enhancer is inactive and the body enhancer is activated. Following our model, the active body enhancer may be disrupting the interaction between insulators placed upstream and downstream of the body enhancer, resulting in the interaction of only two upstream insulators in class II, and the interaction of only two downstream insulators in class III. Therefore, the single downstream insulator which does not interact with the two upstream insulators blocks the body enhancer of class II, while the two downstream insulators which interact with each other allow the body enhancer to bypass to activate the promoter in class III.

Alternatively, regardless of the active or inactive state of the enhancer, three insulators may interact to form distinct chromatin loop domains. Under this assumption, the relative arrangement or position of insulators, enhancers and promoter may lead to

a different topology of chromatin loops, which in turn can differentially affect enhancer-promoter communications. For example, Savitskaya et al showed that an enhancer inside of a loop domain can activate a promoter outside of the loop domain when the distance between two insulators surrounding the enhancer, and therefore the loop formed by the paired insulators, is large enough. As a result, the authors suggested that the formation of the loop domain itself is not a necessary condition to block enhancer activity, and conformational and/or steric hindrances are in fact responsible for blocking the access of the enhancer to the promoter (Savitskaya et al., 2006).

We favor the idea that the active enhancer interrupts the interaction between the upstream and downstream insulators. First, using DNA adenine methyltransferase identification (DamID) method to probe long distance chromatin interactions, Cleard and his colleagues showed that the Fab-7 boundary region interacts with the Abd-B promoter in areas where Abd-B is not expressed (Cleard et al., 2006). On the basis of this result, a molecular switch model was suggested. It states that the boundary elements bind to the Abd-B promoter when Abd-B is in the off-configuration and the boundary elements do not bind to the Abd-B promoter when Abd-B is in the on-configuration (Celniker and Drewell, 2007). In other words, the active or inactive state of *cis*-regulatory regions may determine the on or off-configuration of the molecular switch by interrupting or allowing the interaction between boundary elements and the Abd-B promoter. Second, when the body enhancer was surrounded by two *gypsy* retroviruses, it failed to activate the *yellow* promoter even when the opposite orientation of two *gypsy* retroviruses creates a 15kb loop domain (Figure 1.5B) (Labrador et al.,

2008). This shows that the idea of conformational/steric hindrance is still not enough to explain the functional interaction of chromatin insulators in gene regulation. Nevertheless, the notion that differences in spatial arrangement of enhancers and promoter induced by interactions between multiple insulators might lead to alternative regulatory outputs cannot be ruled out.

The *gypsy* insulator may be involved in gene expression regulation by directly binding to the promoter region.

Several reports have previously suggested that in addition to enhancer-blocking and heterochromatin boundary function, the *gypsy* insulator may also have an effect on the levels of transcription at specific promoters (Golovnin et al., 2005; Soshnev et al., 2008; Wei and Brennan, 2001). We observed that an insulator located upstream of an enhancer can stimulate transcriptional activity when transcription is driven by a tissue specific enhancer, suggesting that insulators may also affect the transcriptional output of genes in a more direct manner than previously thought. Unlike previous observations, our data does not suggest that the insulator itself is sufficient to activate transcription at the *yellow* promoter. It appears, however, as if the insulator is capable of modulating the activity of the enhancer. For example, we could not obtain evidence of transcriptional activation in experiments in which we selectively remove the wing and body enhancers from the regulatory region of the pUAS*y*-λA9 transgenes, leaving an insulator directly upstream of the promoter (pUAS*y*-λA9 Cre2). However, transcription of *yellow* in the wing blades and body cuticle increased significantly when the *gypsy* insulator was

placed upstream of the wing and body enhancers, suggesting that insulators as well as enhancers may actively participate in the process of transcriptional activation.

The traditional insulator properties of enhancer-blocking and heterochromatin boundary function may be interpreted as an indirect manifestation of the effect of upstream insulators on transcription. Eliminating the 1-A2 insulator, the endogenous Su(Hw) binding site found between the *yellow* gene and the *achaete-scute* (*ac*) complex (Golovnin et al., 2003; Parnell et al., 2003), does not result in phenotypes that reflect predictable changes in enhancer-blocking or heterochromatin boundary function at the *y* or *ac* genes (Soshnev et al., 2008). The only phenotype observed in flies carrying a homozygous deletion of the 1-A2 insulator is a significant reduction in the transcription levels of a non-coding RNA located downstream of the insulator. However, the same insulator sequence was shown to have enhancer-blocking activity when tested in an enhancer-blocking assay. These results reinforce our claim that the *gypsy* insulator can modulate transcription levels by demonstrating that it may function as a transcriptional activator also at endogenous sites, but does not discriminate whether the transcriptional activity is independent from the previously described enhancer-blocking and heterochromatin boundary activities.

Insights into the relationship between transcriptional and insulator activities of *gypsy* come from very accurate experiments designed to precisely monitor the transcriptional activity of a number of transgenes flanked by two *gypsy* insulators at different integration sites in the genome (Markstein et al., 2008). Experiments by Markstein et al showed that the boundary activity of *gypsy* insulators actually results in a

20 fold increase in the induced activity of such transgenes, only in the presence of GAL4. The authors suggested from these results that there is not a single locus in the genome where high transgene expression could be reliably induced in every tissue, and the ability of the *gypsy* insulator to modulate transcription is linked to its boundary activity. Thus, the *gypsy* insulators function by preventing the effects of a repressive environment resulting in the boosted expression levels of the transgenes only when flanked by insulators. When transgenes are flanked by *gypsy* insulators, the combined transcriptional activity induced by the *gypsy* insulator and GAL4 boosts transcription to high levels. The interactions between *gypsy* insulators and the GAL4 transcriptional activator shown by Markstein et al. are intriguingly similar to the interactions reported here between upstream *gypsy* insulators and the *yellow* wing and body enhancers. Our data, however, provide evidence that the *yellow* pUAS*y*-λ transgenes have been weakened by the composition of the DNA within the transgene and not by the chromatin environment at the integration site, suggesting that the boost in transcription is due to the presence of the upstream *gypsy* insulator and is independent of position effects. In addition, our data also shows that transcription may be boosted by a single upstream insulator and does not require two insulators flanking both ends of the transgene.

Since single cell measurements using immunofluorescence and fluorescence assisted cell sorting (FACS) have shown that the SV40 enhancer increases the probability that a transfected plasmid would express a reporter gene but does not increase the expression rate in a single cell (Fiering et al., 2000; Walters et al., 1995), it is still controversial whether the enhancer can increase the transcription rate

(progressive model) or the probability of the associated gene to get transcribed ("on or off" model) (Blackwood and Kadonaga, 1998). No matter which is correct for enhancer action, demonstration of direct contacts between long-range enhancers and their target genes using chromosome conformation capture (3C) and RNA TRAP techniques revealed that direct contact between the enhancer and promoter is required for the activation of promoter (Carter et al., 2002; Tolhuis et al., 2002). Moreover, it has been suggested that juxtaposition between the enhancer and promoter precedes direct contact between promoter and enhancer (Polikanov et al., 2007). Therefore, the increase in probability of enhancer-promoter juxtaposition may be intimately related to the increase in transcriptional level. Considering that the enhancers of transgenes integrated into ectopic contexts could not reproduce a normal expression level of endogenous enhancers (Markstein et al., 2008), the juxtaposition of enhancer and promoter may be dependent on neighboring genomic context. In this context, *gypsy* insulators may increase the probability of juxtaposition of the enhancer and promoter or strengthen direct contacts between the enhancer and promoter by interactions between insulators and the transcriptional machinery at promoters. These interactions would also create a potential partition between proximal and distal enhancers that may also explain, in part, their enhancer-blocking activity.

Direct evidence for such interactions is currently lacking but can be supported by the following observations. First, earlier findings demonstrating that the transcriptional coactivator Enhancer of y^2 [E(y)2] interacts with Su(Hw) (Kurshakova et al., 2007) may provide the basis for a hypothetical mechanism in which E(y)2 facilitates interactions

between insulators and promoters. *Drosophila* E(y)2 is a chromatin associated protein that functions as a coactivator on chromatin templates in vitro (Georgieva et al., 2001). In addition, genetic and biochemical experiments have demonstrated interactions between E(y)2 and TAF9, a component of the TFIID general transcription factor. Interestingly E(y)2 is also a homolog of the yeast Sus1 protein, which is a component of the SAGA complex and has a role in transcriptional elongation and nuclear cytoplasm export (Kurshakova et al., 2007). These observations strongly suggest that E(y)2 is most probably a component of the RNA polymerase II transcriptional machinery and open the possibility that E(y)2 may actually facilitate interactions between *gypsy* insulators and components of the transcriptional initiation machinery at the promoter. Second, we observed variegated bristles in all *mod(mdg4)*^{u1} mutant flies containing *gypsy* retrovirus insertions, regardless of position of the *gypsy* insertion (Table 3.1). The *mod(mdg4)*^{u1} mutant line carrying a single *gypsy* retrovirus insertion between the wing enhancer and body enhancer showed black body and variegated bristles, indicating that the body enhancer located closer to the *gypsy* insulator is not affected by silencing, but the bristle enhancer downstream of the promoter is partially silenced (Figure 3.9D). This result can be explained by our speculation that the insulator may interact with the promoter. Their interaction will result in the access of the insulator to the bristle enhancer downstream of the promoter and the local silencing effect, which originated from the insulator in the absence of the Mod(mdg4)67.2 protein, may be transmitted to the bristle enhancer. Third, silencing by the *mod(mdg4)*^{u1} mutation exhibits promoter-specificity (Cai and Levine, 1997; Georgiev and Kozycina, 1996; Georgiev and

Gerasimova, 1989). When the *gypsy* insulator is placed between the enhancer and promoter, in the absence of the Mod(mdg4) protein, the *eve* and *cut* promoters simply show the reversion of the effect of the *gypsy* insertion while the *yellow* and *white* promoters are silenced. Cai et al suggested from these results that the proximal sequences of the *white* and *yellow* promoters are responsible for silencing by the *mod(mdg4)*^{u1} mutation (Cai and Levine, 1997). In other words, this means that insulator proteins may interact with the transcriptional machinery in the promoter region. In addition, Melnikova et al identified that the proximal sequences of *yellow* promoter are located at -100 to -69 from promoter, and also suggested the interaction between these sequences and the *gypsy* insulator (Melnikova et al., 2008).The variegated body phenotypes in the *mod(mdg4)*^{u1} background shown in Figure 3.9B,E,and H indicate that silencing of the *yellow* gene in the absence of Mod(mdg4)67.2 protein is related to changes in chromatin structure such as chromosome condensation, because the variegated pattern is a typical characteristic shown in translocations of the *yellow* gene close to heterochromatin (Gerasimova et al., 1995). ChIP-chip tiling array data from the modENCODE consortium shows that transcription factors encoded by *snr1* (*SNF5-Related-1*) and *brm* (*brahma*) are localized to the *yellow* and *white* promoter region in the pupal stage. These proteins are components of the Brahma (Brm) complex, a yeast SWI/SNF-related chromatin remodeling complex (Dingwall et al., 1995). The ATP-dependent chromatin remodeling complex has been reported to be related to gene activation and repression through collaboration with gene-specific transcription factors and histone modifications (Kingston et al., 1996; Marenda et al., 2004; Peterson, 2002).

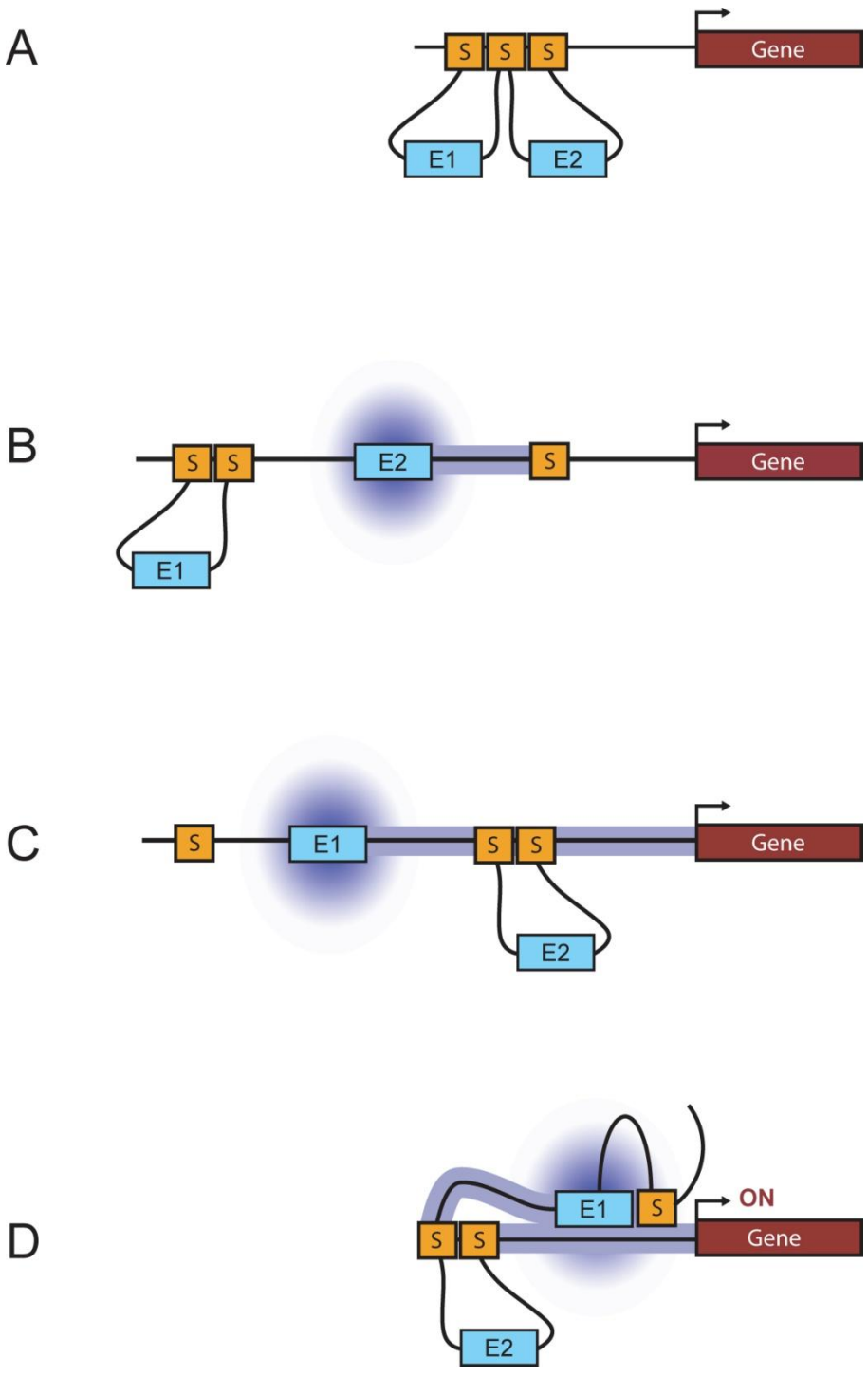
Interestingly, the combined mutations of *mod(mdg4)* and *brm* genes significantly increased the frequency of transformations, such as the change of haltere to wing, indicating that a genetic interaction exists between them (Gerasimova and Corces, 1998). Also, Gerasimova and Corces showed by genetic analysis that several other polycomb group and trithorax group proteins are involved in the function of the *gypsy* insulator. Nevertheless, these proteins do not co-localize with endogenous *gypsy* insulators on polytene chromosomes (Gerasimova and Corces, 1998) and the genome wide mapping of *Drosophila* insulator protein binding sites revealed that most Su(Hw) bindings sites are located far from the promoter (Bushey et al., 2009; Negre et al., 2010). This suggests that the interaction between insulator proteins and polycomb group and/or tirthorax group proteins is achieved at a long-distance, accompanying the formation of chromatin loop domains, and the interaction is reversible rather than tightly bound.

Conclusions and Remarks

The current data presented here suggests a model in which *gypsy* insulators as well as endogenous *gypsy* insulators may function by directly regulating gene expression in the genome. When enhancers are inactive, insulators interact with themselves generating loop domains that directly place enhancers inside of the domain (Figure 4.1A). Enhancers activated at a specific time during development or tissue specific differentiation function as a molecular switch disrupting the interactions between insulators flanking the enhancer. As a result, enhancers are released from the loop

Figure 4.1 Schematic models explaining the regulation of gene expression by *gypsy* insulators.

Blue boxes with E1and E2 indicate enhancers and *gypsy* insulators are depicted as orange boxes with S. The cognate gene of enhancer 2(E2) is illustrated as red box and broken arrow marks promoter region. Activation and chromatin tracking of enhancer are shown as halo and the violet colored thick line, respectively. **(A)** When two enhancers (E1 and E2) are inactive, three insulators interact with themselves, generating two chromatin loops. **(B)** When enhancer 2 (E2) is activated, active E2 enhancer disrupts the interaction between upstream and downstream insulators. The tracking of E2 enhancer along the chromatin fiber is blocked by the downstream insulator, and so the E2 enhancer cannot activate the promoter. **(C)** When enhancer 1 (E1) is activated, the upstream insulator is dissociated but the two insulators downstream still interact. Tracking of E1 enhancer along chromatin fiber bypasses the paired downstream insulators to reach the promoter. **(D)** Subsequently or concurrently with tracking, a loop domain is formed by the interaction of E1 enhancer and promoter. The binding of upstream insulator to a proximal region of the promoter facilitates gene expression by the E1 enhancer.


A
S
S
S
Gene
E1
E2
B
S
S
E2
S
Gene
E1
C
S
E1
S
S
Gene
E2
D
E1
S
ON
S
S
Gene
E2

domain and only interactions between two pairwise insulators remains active (Figure 4.1B and C). Experiments directly measuring changes in the intervening chromatin after induction of transcription, have revealed evidence suggesting that enhancers may track the chromatin fiber by acetylating intervening histones (Hatzis and Talianidis, 2002). In a similar experiment, Wang et al. demonstrated that a phosphorylated form of the RNA polymerase II tracks along the chromatin fiber from the enhancer to the promoter after induction of transcription of the Prostate Specific Antigen [PSA] and upon binding of the Androgen Receptor to an enhancer located 4 kb upstream of the PSA promoter (Wang et al., 2005). Therefore, active enhancers track the chromatin fiber to find their cognate promoters. In a configuration like the one shown in Figure 4.1B, chromatin tracking of enhancer 2 will be blocked by the downstream single insulator and activation of the promoter will not be induced by enhancer 2. In the configuration like the one shown in Figure 4.1C, chromatin tracking of active enhancer 1 toward the promoter will be dependent on the relative orientation of two downstream insulators. Once paired insulators allow tracking of the enhancer to bypass the enhancer-blocking property of the insulator, juxtaposition of enhancer 1 with the promoter will cause the formation of a chromatin loop. The interaction between the insulator upstream of enhancer 1 and the promoter proximal region may stabilize or strengthen the direct contact between enhancer 1 and the promoter, probably through the interaction between components of the insulator protein complex and the transcriptional machinery or chromatin remodeling complexes associated to the promoters. This will result in the induction of strong gene expression levels by enhancer 1 (Figure 4.1D).

Results presented here open a new avenue toward the study of the molecular mechanisms that enhancers use to activate remote promoters as well as the mechanism that insulators use to control the activity of enhancers. Our results also provide new approaches to understand genome evolution as well as the molecular basis of genome organization and its role in gene expression regulation.

LIST OF REFERENCES

Adryan, B., Woerfel, G., Birch-Machin, I., Gao, S., Quick, M., Meadows, L., Russell, S., and White, R. (2007). Genomic mapping of Suppressor of Hairy-wing binding sites in *Drosophila*. Genome Biol *8*, R167.

Ahmad, K.F., Engel, C.K., and Prive, G.G. (1998). Crystal structure of the BTB domain from PLZF. Proc Natl Acad Sci U S A *95*, 12123-12128.

Akbari, O.S., Bousum, A., Bae, E., and Drewell, R.A. (2006). Unraveling cis-regulatory mechanisms at the abdominal-A and Abdominal-B genes in the *Drosophila* bithorax complex. Dev Biol *293*, 294-304.

Alexandre, C., Jacinto, A., and Ingham, P.W. (1996). Transcriptional activation of hedgehog target genes in *Drosophila* is mediated directly by the cubitus interruptus protein, a member of the GLI family of zinc finger DNA-binding proteins. Genes Dev *10*, 2003-2013.

Andrulis, E.D., Neiman, A.M., Zappulla, D.C., and Sternglanz, R. (1998). Perinuclear localization of chromatin facilitates transcriptional silencing. Nature *394*, 592-595.

Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482-485.

Bell, A.C., West, A.G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell *98*, 387-396.

Belozerov, V.E., Majumder, P., Shen, P., and Cai, H.N. (2003). A novel boundary element may facilitate independent gene regulation in the Antennapedia complex of *Drosophila*. EMBO J *22*, 3113-3121.

Blackwood, E.M., and Kadonaga, J.T. (1998). Going the distance: a current view of enhancer action. Science *281*, 60-63.

Blanton, J., Gaszner, M., and Schedl, P. (2003). Protein:protein interactions and the pairing of boundary elements in vivo. Genes Dev *17*, 664-675.

Bode, J., Goetze, S., Heng, H., Krawetz, S.A., and Benham, C. (2003). From DNA structure to gene expression: mediators of nuclear compartmentalization and dynamics. Chromosome Res *11*, 435-445.

Bode, J., Schlake, T., Rios-Ramirez, M., Mielke, C., Stengert, M., Kay, V., and Klehr-Wirth, D. (1995). Scaffold/matrix-attached regions: structural properties creating transcriptionally active loci. Int Rev Cytol *162A*, 389-454.

Borden, J., and Manuelidis, L. (1988). Movement of the X chromosome in epilepsy. Science *242*, 1687-1691.

Brasset, E., Bantignies, F., Court, F., Cheresiz, S., Conte, C., and Vaury, C. (2007). Idefix insulator activity can be modulated by nearby regulatory elements. Nucleic Acids Res *35*, 2661-2670.

Brasset, E., and Vaury, C. (2005). Insulators are fundamental components of the eukaryotic genomes. Heredity *94*, 571-576.

Buchner, K., Roth, P., Schotta, G., Krauss, V., Saumweber, H., Reuter, G., and Dorn, R. (2000). Genetic and molecular complexity of the position effect variegation modifier mod(mdg4) in *Drosophila*. Genetics *155*, 141-157.

Bushey, A.M., Dorman, E.R., and Corces, V.G. (2008). Chromatin insulators: regulatory mechanisms and epigenetic inheritance. Molecular cell *32*, 1-9.

Bushey, A.M., Ramos, E., and Corces, V.G. (2009). Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. Genes & development *23*, 1338-1350.

Butcher, R.D., Chodagam, S., Basto, R., Wakefield, J.G., Henderson, D.S., Raff, J.W., and Whitfield, W.G. (2004). The *Drosophila* centrosome-associated protein CP190 is essential for viability but not for cell division. J Cell Sci *117*, 1191-1199.

Butler, J.E., and Kadonaga, J.T. (2001). Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. Genes Dev *15*, 2515-2519.

Byrd, K., and Corces, V.G. (2003). Visualization of chromatin domains created by the *gypsy* insulator of *Drosophila*. J Cell Biol *162*, 565-574.

Cai, H., and Levine, M. (1995). Modulation of enhancer-promoter interactions by insulators in the *Drosophila* embryo. Nature *376*, 533-536.

Cai, H.N., and Levine, M. (1997). The *gypsy* insulator can function as a promoter-specific silencer in the *Drosophila* embryo. EMBO J *16*, 1732-1741.

Cai, H.N., and Shen, P. (2001). Effects of cis arrangement of chromatin insulators on enhancer-blocking activity. Science *291*, 493-495.

Capelson, M., and Corces, V.G. (2005). The ubiquitin ligase dTopors directs the nuclear organization of a chromatin insulator. Mol Cell *20*, 105-116.

Capelson, M., and Corces, V.G. (2006). SUMO conjugation attenuates the activity of the *gypsy* chromatin insulator. EMBO J *25*, 1906-1914.

Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. Nat Genet *32*, 623-626.

Casolari, J.M., Brown, C.R., Komili, S., West, J., Hieronymus, H., and Silver, P.A. (2004). Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. Cell *117*, 427-439.

Celniker, S.E., Dillon, L.A., Gerstein, M.B., Gunsalus, K.C., Henikoff, S., Karpen, G.H., Kellis, M., Lai, E.C., Lieb, J.D., MacAlpine, D.M.*, et al.* (2009). Unlocking the secrets of the genome. Nature *459*, 927-930.

Celniker, S.E., and Drewell, R.A. (2007). Chromatin looping mediates boundary element promoter interactions. Bioessays *29*, 7-10.

Choi, Y.J., Lee, G., Hall, J.C., and Park, J.H. (2005). Comparative analysis of Corazonin-encoding genes (Crz's) in *Drosophila* species and functional insights into Crz-expressing neurons. J Comp Neurol *482*, 372-385.

Cleard, F., Moshkin, Y., Karch, F., and Maeda, R.K. (2006). Probing long-distance regulatory interactions in the *Drosophila melanogaster* bithorax complex using Dam identification. Nat Genet *38*, 931-935.

Cockerill, P.N., and Garrard, W.T. (1986). Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. Cell *44*, 273-282.

Comet, I., Savitskaya, E., Schuettengruber, B., Negre, N., Lavrov, S., Parshikov, A., Juge, F., Gracheva, E., Georgiev, P., and Cavalli, G. (2006). PRE-mediated bypass of two Su(Hw) insulators targets PcG proteins to a downstream promoter. Dev Cell *11*, 117-124.

Coney, L.R., and Roeder, G.S. (1988). Control of yeast gene expression by transposable elements: maximum expression requires a functional Ty activator sequence and a defective Ty promoter. Mol Cell Biol *8*, 4009-4017.

Conte, C., Dastugue, B., and Vaury, C. (2002a). Coupling of enhancer and insulator properties identified in two retrotransposons modulates their mutagenic impact on nearby genes. Mol Cell Biol *22*, 1767-1777.

Conte, C., Dastugue, B., and Vaury, C. (2002b). Promoter competition as a mechanism of transcriptional interference mediated by retrotransposons. EMBO J *21*, 3908-3916.

Cremer, T., Cremer, M., Dietzel, S., Muller, S., Solovei, I., and Fakan, S. (2006). Chromosome territories--a functional nuclear landscape. Curr Opin Cell Biol *18*, 307-316.

Cremer, T., Lichter, P., Borden, J., Ward, D.C., and Manuelidis, L. (1988). Detection of chromosome aberrations in metaphase and interphase tumor cells by in situ hybridization using chromosome-specific library probes. Hum Genet *80*, 235-246.

Croston, G.E., and Kadonaga, J.T. (1993). Role of chromatin structure in the regulation of transcription by RNA polymerase II. Curr Opin Cell Biol *5*, 417-423.

Csink, A.K., Linsk, R., and Birchler, J.A. (1994). The Lighten up (Lip) gene of *Drosophila melanogaster*, a modifier of retroelement expression, position effect variegation and white locus insertion alleles. Genetics *138*, 153-163.

de Bruin, D., Zaman, Z., Liberatore, R.A., and Ptashne, M. (2001). Telomere looping permits gene activation by a downstream UAS in yeast. Nature *409*, 109-113.

de Wit, E., Greil, F., and van Steensel, B. (2005). Genome-wide HP1 binding in *Drosophila*: developmental plasticity and genomic targeting signals. Genome Res *15*, 1265-1273.

Dej, K.J., Gerasimova, T., Corces, V.G., and Boeke, J.D. (1998). A hotspot for the *Drosophila gypsy* retroelement in the ovo locus. Nucleic Acids Res *26*, 4019-4025.

Dekker, J. (2008). Gene regulation in the third dimension. Science *319*, 1793-1794.

Dingwall, A.K., Beek, S.J., McCallum, C.M., Tamkun, J.W., Kalpana, G.V., Goff, S.P., and Scott, M.P. (1995). The *Drosophila* snr1 and brm proteins are related to yeast SWI/SNF proteins and are components of a large protein complex. Mol Biol Cell *6*, 777-791.

Dohmen, R.J. (2004). SUMO protein modification. Biochim Biophys Acta *1695*, 113-131.

Dong, S., Zhu, J., Reid, A., Strutt, P., Guidez, F., Zhong, H.J., Wang, Z.Y., Licht, J., Waxman, S., Chomienne, C.*, et al.* (1996). Amino-terminal protein-protein interaction motif (POZ-domain) is responsible for activities of the promyelocytic leukemia zinc finger-retinoic acid receptor-alpha fusion protein. Proc Natl Acad Sci U S A *93*, 3624-3629.

Donze, D., and Kamakaka, R.T. (2001). RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in Saccharomyces cerevisiae. EMBO J *20*, 520-531.

Eaton, S., and Kornberg, T.B. (1990). Repression of ci-D in posterior compartments of *Drosophila* by engrailed. Genes Dev *4*, 1068-1077.

Eberl, D.F., Duyf, B.J., and Hilliker, A.J. (1993). The role of heterochromatin in the expression of a heterochromatic gene, the rolled locus of *Drosophila melanogaster*. Genetics *134*, 277-292.

Espinas, M.L., Jimenez-Garcia, E., Vaquero, A., Canudas, S., Bernues, J., and Azorin, F. (1999). The N-terminal POZ domain of GAGA mediates the formation of oligomers that bind DNA with high affinity and specificity. J Biol Chem *274*, 16461-16469.

Fiering, S., Whitelaw, E., and Martin, D.I. (2000). To be or not to be active: the stochastic nature of enhancer action. Bioessays *22*, 381-387.

Fischer, T., Rodriguez-Navarro, S., Pereira, G., Racz, A., Schiebel, E., and Hurt, E. (2004). Yeast centrin Cdc31 is linked to the nuclear mRNA export machinery. Nat Cell Biol *6*, 840-848.

Fourel, G., Revardel, E., Koering, C.E., and Gilson, E. (1999). Cohabitation of insulators and silencing elements in yeast subtelomeric regions. EMBO J *18*, 2522-2537.

Galande, S., Purbey, P.K., Notani, D., and Kumar, P.P. (2007). The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1. Curr Opin Genet Dev *17*, 408-414.

Gasser, S.M., and Laemmli, U.K. (1986). Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of D. melanogaster. Cell *46*, 521-530.

Gaszner, M., Vazquez, J., and Schedl, P. (1999). The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction. Genes Dev *13*, 2098-2107.

Gause, M., Morcillo, P., and Dorsett, D. (2001). Insulation of enhancer-promoter communication by a *gypsy* transposon insert in the *Drosophila* cut gene: cooperation between suppressor of hairy-wing and modifier of mdg4 proteins. Mol Cell Biol *21*, 4807-4817.

Gdula, D.A., and Corces, V.G. (1997). Characterization of functional domains of the su(Hw) protein that mediate the silencing effect of mod(mdg4) mutations. Genetics *145*, 153-161.

Georgiev, P., and Kozycina, M. (1996). Interaction between mutations in the suppressor of Hairy wing and modifier of mdg4 genes of *Drosophila melanogaster* affecting the phenotype of *gypsy*-induced mutations. Genetics *142*, 425-436.

Georgiev, P.G., and Gerasimova, T.I. (1989). Novel genes influencing the expression of the yellow locus and mdg4 (*gypsy*) in *Drosophila melanogaster*. Mol Gen Genet *220*, 121-126.

Georgieva, S., Nabirochkina, E., Dilworth, F.J., Eickhoff, H., Becker, P., Tora, L., Georgiev, P., and Soldatov, A. (2001). The novel transcription factor e(y)2 interacts with TAF(II)40 and potentiates transcription activation on chromatin templates. Mol Cell Biol *21*, 5223-5231.

Gerasimova, T.I., Byrd, K., and Corces, V.G. (2000). A chromatin insulator determines the nuclear localization of DNA. Mol Cell *6*, 1025-1035.

Gerasimova, T.I., and Corces, V.G. (1996). Boundary and insulator elements in chromosomes. Curr Opin Genet Dev *6*, 185-192.

Gerasimova, T.I., and Corces, V.G. (1998). Polycomb and trithorax group proteins mediate the function of a chromatin insulator. Cell *92*, 511-521.

Gerasimova, T.I., Gdula, D.A., Gerasimov, D.V., Simonova, O., and Corces, V.G. (1995). A *Drosophila* protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. Cell *82*, 587-597.

Geyer, P.K., and Clark, I. (2002). Protecting against promiscuity: the regulatory role of insulators. Cell Mol Life Sci *59*, 2112-2127.

Geyer, P.K., Spana, C., and Corces, V.G. (1986). On the molecular mechanism of *gypsy*-induced mutations at the yellow locus of *Drosophila melanogaster*. EMBO J *5*, 2657-2662.

Ghosh, D., Gerasimova, T.I., and Corces, V.G. (2001). Interactions between the Su(Hw) and Mod(mdg4) proteins required for *gypsy* insulator function. EMBO J *20*, 2518-2527.

Golovnin, A., Biryukova, I., Romanova, O., Silicheva, M., Parshikov, A., Savitskaya, E., Pirrotta, V., and Georgiev, P. (2003). An endogenous Su(Hw) insulator separates the

yellow gene from the Achaete-scute gene complex in *Drosophila*. Development (Cambridge, England) *130*, 3249-3258.

Golovnin, A., Melnick, E., Mazur, A., and Georgiev, P. (2005). *Drosophila* Su(Hw) insulator can stimulate transcription of a weakened yellow promoter over a distance. Genetics *170*, 1133-1142.

Golovnin, A., Melnikova, L., Volkov, I., Kostuchenko, M., Galkin, A.V., and Georgiev, P. (2008). 'Insulator bodies' are aggregates of proteins but not of insulators. EMBO Rep *9*, 440-445.

Gomos-Klein, J., Harrow, F., Alarcon, J., and Ortiz, B.D. (2007). CTCF-independent, but not CTCF-dependent, elements significantly contribute to TCR-alpha locus control region activity. J Immunol *179*, 1088-1095.

Grewal, S.I., and Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. Science *301*, 798-802.

Grimm, S., and Pflugfelder, G.O. (1996). Control of the gene optomotor-blind in *Drosophila* wing development by decapentaplegic and wingless. Science *271*, 1601-1604.

Guarente, L., Lalonde, B., Gifford, P., and Alani, E. (1984). Distinctly regulated tandem upstream activation sites mediate catabolite repression of the CYC1 gene of S. cerevisiae. Cell *36*, 503-511.

Hagstrom, K., Muller, M., and Schedl, P. (1996). Fab-7 functions as a chromatin domain boundary to ensure proper segment specification by the *Drosophila* bithorax complex. Genes Dev *10*, 3202-3215.

Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature *405*, 486-489.

Harrison, D.A., Gdula, D.A., Coyne, R.S., and Corces, V.G. (1993). A leucine zipper domain of the suppressor of Hairy-wing protein mediates its repressive effect on enhancer function. Genes Dev *7*, 1966-1978.

Hatzis, P., and Talianidis, I. (2002). Dynamics of enhancer-promoter communication during differentiation-induced gene activation. Mol Cell *10*, 1467-1477.

Hay, R.T. (2005). SUMO: a history of modification. Mol Cell *18*, 1-12.

Heger, P., Marin, B., and Schierenberg, E. (2009). Loss of the insulator protein CTCF during nematode evolution. BMC molecular biology *10*, 84.

Hepker, J., Wang, Q.T., Motzny, C.K., Holmgren, R., and Orenic, T.V. (1997). *Drosophila* cubitus interruptus forms a negative feedback loop with patched and regulates expression of Hedgehog target genes. Development *124*, 549-558.

Holohan, E.E., Kwong, C., Adryan, B., Bartkuhn, M., Herold, M., Renkawitz, R., Russell, S., and White, R. (2007). CTCF genomic binding sites in *Drosophila* and the organisation of the bithorax complex. PLoS Genet *3*, e112.

Hou, C., Zhao, H., Tanimoto, K., and Dean, A. (2008). CTCF-dependent enhancer-blocking by alternative chromatin loop formation. Proceedings of the National Academy of Sciences of the United States of America *105*, 20398-20403.

Ishizuka, A., Siomi, M.C., and Siomi, H. (2002). A *Drosophila* fragile X protein interacts with components of RNAi and ribosomal proteins. Genes Dev *16*, 2497-2508.

Jimenez, M., and Goday, C. (1993). A centrosome-associated antibody from *Drosophila melanogaster* reveals a new microtubule-dependent structure in the equatorial zone of Parascaris univalens embryos. J Cell Sci *106 (Pt 3)*, 719-730.

Kalos, M., and Fournier, R.E. (1995). Position-independent transgene expression mediated by boundary elements from the apolipoprotein B chromatin domain. Mol Cell Biol *15*, 198-207.

Karch, F., Galloni, M., Sipos, L., Gausz, J., Gyurkovics, H., and Schedl, P. (1994). Mcp and Fab-7: molecular analysis of putative boundaries of cis-regulatory domains in the bithorax complex of *Drosophila melanogaster*. Nucleic Acids Res *22*, 3138-3146.

Karch, F., Weiffenbach, B., Peifer, M., Bender, W., Duncan, I., Celniker, S., Crosby, M., and Lewis, E.B. (1985). The abdominal region of the bithorax complex. Cell *43*, 81-96.

Kellum, R., and Schedl, P. (1992). A group of scs elements function as domain boundaries in an enhancer-blocking assay. Mol Cell Biol *12*, 2424-2431.

Kiefer, C.M., Hou, C., Little, J.A., and Dean, A. (2008). Epigenetics of beta-globin gene regulation. Mutat Res *647*, 68-76.

Kim, J., Shen, B., Rosen, C., and Dorsett, D. (1996). The DNA-binding and enhancer-blocking domains of the *Drosophila* suppressor of Hairy-wing protein. Mol Cell Biol *16*, 3381-3392.

Kingston, R.E., Bunker, C.A., and Imbalzano, A.N. (1996). Repression and activation by multiprotein complexes that alter chromatin structure. Genes Dev *10*, 905-920.

Kleinjan, D.A., and van Heyningen, V. (2005). Long-range control of gene expression: emerging mechanisms and disruption in disease. Am J Hum Genet *76*, 8-32.

Kostyuchenko, M.V., Savitskaya, E.E., Volkov, I.A., Golovnin, A.K., and Georgiev, P.G. (2008). Study of functional interaction between three copies of the insulator from the MDG4 transposable element in the model system of the miniwhite gene of *Drosophila melanogaster*. Dokl Biochem Biophys *421*, 239-243.

Kuhn-Parnell, E.J., Helou, C., Marion, D.J., Gilmore, B.L., Parnell, T.J., Wold, M.S., and Geyer, P.K. (2008). Investigation of the properties of non-*gypsy* suppressor of hairy-wing-binding sites. Genetics *179*, 1263-1273.

Kuhn, E.J., and Geyer, P.K. (2003). Genomic insulators: connecting properties to mechanism. Curr Opin Cell Biol *15*, 259-265.

Kuhn, E.J., Viering, M.M., Rhodes, K.M., and Geyer, P.K. (2003). A test of insulator interactions in *Drosophila*. EMBO J *22*, 2463-2471.

Kurshakova, M., Maksimenko, O., Golovnin, A., Pulina, M., Georgieva, S., Georgiev, P., and Krasnov, A. (2007). Evolutionarily conserved E(y)2/Sus1 protein is essential for the barrier activity of Su(Hw)-dependent insulators in *Drosophila*. Molecular cell *27*, 332-338.

Kyrchanova, O., Chetverina, D., Maksimenko, O., Kullyev, A., and Georgiev, P. (2008). Orientation-dependent interaction between *Drosophila* insulators is a property of this class of regulatory elements. Nucleic Acids Res *36*, 7019-7028.

Kyrchanova, O., Toshchakov, S., Parshikov, A., and Georgiev, P. (2007). Study of the functional interaction between Mcp insulators from the *Drosophila* bithorax complex: effects of insulator pairing on enhancer-promoter communication. Mol Cell Biol *27*, 3035-3043.

Labrador, M., and Corces, V.G. (2001). Protein determinants of insertional specificity for the *Drosophila gypsy* retrovirus. Genetics *158*, 1101-1110.

Labrador, M., and Corces, V.G. (2002). Setting the boundaries of chromatin domains and nuclear organization. Cell *111*, 151-154.

Labrador, M., Mongelard, F., Plata-Rengifo, P., Baxter, E.M., Corces, V.G., and Gerasimova, T.I. (2001). Protein encoding by both DNA strands. Nature *409*, 1000.

Labrador, M., Sha, K., Li, A., and Corces, V.G. (2008). Insulator and Ovo proteins determine the frequency and specificity of insertion of the *gypsy* retrotransposon in *Drosophila melanogaster*. Genetics *180*, 1367-1378.

Lagarkova, M.A., Svetlova, E., Giacca, M., Falaschi, A., and Razin, S.V. (1998). DNA loop anchorage region colocalizes with the replication origin located downstream to the human gene encoding lamin B2. J Cell Biochem *69*, 13-18.

Lei, E.P., and Corces, V.G. (2006). RNA interference machinery influences the nuclear organization of a chromatin insulator. Nat Genet *38*, 936-941.

Lettice, L.A., Heaney, S.J., Purdie, L.A., Li, L., de Beer, P., Oostra, B.A., Goode, D., Elgar, G., Hill, R.E., and de Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Hum Mol Genet *12*, 1725-1735.

Ling, J.Q., Li, T., Hu, J.F., Vu, T.H., Chen, H.L., Qiu, X.W., Cherry, A.M., and Hoffman, A.R. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science *312*, 269-272.

Luo, R.X., and Dean, D.C. (1999). Chromatin remodeling and transcriptional regulation. J Natl Cancer Inst *91*, 1288-1294.

Mahmoudi, T., Katsani, K.R., and Verrijzer, C.P. (2002). GAGA can mediate enhancer function in trans by linking two separate DNA molecules. EMBO J *21*, 1775-1781.

Majumder, P., and Cai, H.N. (2003). The functional analysis of insulator interactions in the *Drosophila* embryo. Proc Natl Acad Sci U S A *100*, 5223-5228.

Maksimenko, O., Golovnin, A., and Georgiev, P. (2008). Enhancer-promoter communication is regulated by insulator pairing in a *Drosophila* model bigenic locus. Molecular and cellular biology *28*, 5469-5477.

Marenda, D.R., Zraly, C.B., and Dingwall, A.K. (2004). The *Drosophila* Brahma (SWI/SNF) chromatin remodeling complex exhibits cell-type specific activation and repression functions. Dev Biol *267*, 279-293.

Markstein, M., Pitsouli, C., Villalta, C., Celniker, S.E., and Perrimon, N. (2008). Exploiting position effects and the *gypsy* retrovirus insulator to engineer precisely expressed transgenes. Nature genetics *40*, 476-483.

Marshall, W.F., Straight, A., Marko, J.F., Swedlow, J., Dernburg, A., Belmont, A., Murray, A.W., Agard, D.A., and Sedat, J.W. (1997). Interphase chromosomes undergo constrained diffusional motion in living cells. Curr Biol *7*, 930-939.

Martens, J.H., O'Sullivan, R.J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P., and Jenuwein, T. (2005). The profile of repeat-associated histone lysine methylation states in the mouse epigenome. EMBO J *24*, 800-812.

Melnikova, L., Kostuchenko, M., Silicheva, M., and Georgiev, P. (2008). *Drosophila gypsy* insulator and yellow enhancers regulate activity of yellow promoter through the same regulatory element. Chromosoma *117*, 137-145.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P.*, et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Mirkovitch, J., Mirault, M.E., and Laemmli, U.K. (1984). Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. Cell *39*, 223-232.

Mirsky, A.E., and Allfrey, V. (1960). Biochemical activities of the cell nucleus. Dis Nerv Syst *21(2)Suppl*, 23-28.

Mohan, M., Bartkuhn, M., Herold, M., Philippen, A., Heinl, N., Bardenhagen, I., Leers, J., White, R.A., Renkawitz-Pohl, R., Saumweber, H.*, et al.* (2007). The *Drosophila* insulator proteins CTCF and CP190 link enhancer blocking to body patterning. EMBO J *26*, 4203-4214.

Mongelard, F., and Corces, V.G. (2001). Two insulators are not better than one. Nature structural biology *8*, 192-194.

Mongelard, F., Labrador, M., Baxter, E.M., Gerasimova, T.I., and Corces, V.G. (2002). Trans-splicing as a novel mechanism to explain interallelic complementation in *Drosophila*. Genetics *160*, 1481-1487.

Moon, H., Filippova, G., Loukinov, D., Pugacheva, E., Chen, Q., Smith, S.T., Munhall, A., Grewe, B., Bartkuhn, M., Arnold, R.*, et al.* (2005). CTCF is conserved from *Drosophila* to humans and confers enhancer blocking of the Fab-8 insulator. EMBO Rep *6*, 165-170.

Muller, J. (2000). Transcriptional control: The benefits of selective insulation. Curr Biol *10*, R241-244.

Muravyova, E., Golovnin, A., Gracheva, E., Parshikov, A., Belenkaya, T., Pirrotta, V., and Georgiev, P. (2001). Loss of insulator activity by paired Su(Hw) chromatin insulators. Science *291*, 495-498.

Nabirochkin, S., Ossokina, M., and Heidmann, T. (1998). A nuclear matrix/scaffold attachment region co-localizes with the *gypsy* retrotransposon insulator sequence. J Biol Chem *273*, 2473-2479.

Negre, N., Brown, C.D., Shah, P.K., Kheradpour, P., Morrison, C.A., Henikoff, J.G., Feng, X., Ahmad, K., Russell, S., White, R.A.*, et al.* (2010). A comprehensive map of insulator elements for the *Drosophila* genome. PLoS Genet *6*, e1000814.

Nobrega, M.A., Ovcharenko, I., Afzal, V., and Rubin, E.M. (2003). Scanning human gene deserts for long-range enhancers. Science *302*, 413.

Oegema, K., Whitfield, W.G., and Alberts, B. (1995). The cell cycle-dependent localization of the CP190 centrosomal protein is determined by the coordinate action of two separable domains. J Cell Biol *131*, 1261-1273.

Ohtsuki, S., and Levine, M. (1998). GAGA mediates the enhancer blocking activity of the eve promoter in the *Drosophila* embryo. Genes Dev *12*, 3325-3330.

Pai, C.Y., Lei, E.P., Ghosh, D., and Corces, V.G. (2004). The centrosomal protein CP190 is a component of the *gypsy* chromatin insulator. Mol Cell *16*, 737-748.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T.*, et al.* (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell *132*, 422-433.

Parnell, T.J., and Geyer, P.K. (2000). Differences in insulator properties revealed by enhancer blocking assays on episomes. EMBO J *19*, 5864-5874.

Parnell, T.J., Kuhn, E.J., Gilmore, B.L., Helou, C., Wold, M.S., and Geyer, P.K. (2006). Identification of genomic sites that bind the *Drosophila* suppressor of Hairy-wing insulator protein. Mol Cell Biol *26*, 5983-5993.

Parnell, T.J., Viering, M.M., Skjesol, A., Helou, C., Kuhn, E.J., and Geyer, P.K. (2003). An endogenous suppressor of hairy-wing insulator separates regulatory domains in *Drosophila*. Proc Natl Acad Sci U S A *100*, 13436-13441.

Peterson, C.L. (2002). Chromatin remodeling enzymes: taming the machines. Third in review series on chromatin dynamics. EMBO Rep *3*, 319-322.

Petrov, A., Pirozhkova, I., Carnac, G., Laoudj, D., Lipinski, M., and Vassetzky, Y.S. (2006). Chromatin loop domain organization within the 4q35 locus in

facioscapulohumeral dystrophy patients versus normal human myoblasts. Proc Natl Acad Sci U S A *103*, 6982-6987.

Pflugfelder, G.O., Roth, H., and Poeck, B. (1992). A homology domain shared between *Drosophila* optomotor-blind and mouse Brachyury is involved in DNA binding. Biochem Biophys Res Commun *186*, 918-925.

Phi-Van, L., and Stratling, W.H. (1996). Dissection of the ability of the chicken lysozyme gene 5' matrix attachment region to stimulate transgene expression and to dampen position effects. Biochemistry *35*, 10735-10742.

Pikaart, M.J., Recillas-Targa, F., and Felsenfeld, G. (1998). Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. Genes Dev *12*, 2852-2862.

Pinkel, D., Landegent, J., Collins, C., Fuscoe, J., Segraves, R., Lucas, J., and Gray, J. (1988). Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. Proc Natl Acad Sci U S A *85*, 9138-9142.

Polikanov, Y.S., Bondarenko, V.A., Tchernaenko, V., Jiang, Y.I., Lutter, L.C., Vologodskii, A., and Studitsky, V.M. (2007). Probability of the site juxtaposition determines the rate of protein-mediated DNA looping. Biophys J *93*, 2726-2731.

Prioleau, M.N., Nony, P., Simpson, M., and Felsenfeld, G. (1999). An insulator element and condensed chromatin region separate the chicken beta-globin locus from an independently regulated erythroid-specific folate receptor gene. EMBO J *18*, 4035-4048.

Prud'homme, N., Gans, M., Masson, M., Terzian, C., and Bucheton, A. (1995). Flamenco, a gene controlling the *gypsy* retrovirus of *Drosophila melanogaster*. Genetics *139*, 697-711.

Ramos, E., Ghosh, D., Baxter, E., and Corces, V.G. (2006). Genomic organization of *gypsy* chromatin insulators in *Drosophila melanogaster*. Genetics *172*, 2337-2349.

Razin, S.V., Iarovaia, O.V., Sjakste, N., Sjakste, T., Bagdoniene, L., Rynditch, A.V., Eivazova, E.R., Lipinski, M., and Vassetzky, Y.S. (2007). Chromatin domains and regulation of transcription. J Mol Biol *369*, 597-607.

Razin, S.V., Kekelidze, M.G., Lukanidin, E.M., Scherrer, K., and Georgiev, G.P. (1986). Replication origins are attached to the nuclear skeleton. Nucleic Acids Res *14*, 8189-8207.

Riddle, N.C., and Elgin, S.C. (2006). The dot chromosome of *Drosophila*: insights into chromatin states and their change over evolutionary time. Chromosome Res *14*, 405-416.

Rodin, S., and Georgiev, P. (2005). Handling three regulatory elements in one transgene: combined use of cre-lox, FLP-FRT, and I-SceI recombination systems. Biotechniques *39*, 871-876.

Rodriguez-Navarro, S., Fischer, T., Luo, M.J., Antunez, O., Brettschneider, S., Lechner, J., Perez-Ortin, J.E., Reed, R., and Hurt, E. (2004). Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. Cell *116*, 75-86.

Roseman, R.R., Pirrotta, V., and Geyer, P.K. (1993). The su(Hw) protein insulates expression of the *Drosophila melanogaster* white gene from chromosomal position-effects. EMBO J *12*, 435-442.

Rubio, E.D., Reiss, D.J., Welcsh, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A *105*, 8309-8314.

Sarot, E., Payen-Groschene, G., Bucheton, A., and Pelisson, A. (2004). Evidence for a piwi-dependent RNA silencing of the *gypsy* endogenous retrovirus by the *Drosophila melanogaster* flamenco gene. Genetics *166*, 1313-1321.

Savitskaya, E., Melnikova, L., Kostuchenko, M., Kravchenko, E., Pomerantseva, E., Boikova, T., Chetverina, D., Parshikov, A., Zobacheva, P., Gracheva, E*., et al.* (2006). Study of long-distance functional interactions between Su(Hw) insulators that can regulate enhancer-promoter communication in *Drosophila melanogaster*. Mol Cell Biol *26*, 754-761.

Schwartz, C., Locke, J., Nishida, C., and Kornberg, T.B. (1995). Analysis of cubitus interruptus regulation in *Drosophila* embryos and imaginal disks. Development *121*, 1625-1635.

Scott, K.C., Taubman, A.D., and Geyer, P.K. (1999). Enhancer blocking by the *Drosophila gypsy* insulator depends upon insulator anatomy and enhancer strength. Genetics *153*, 787-798.

Secombe, J., and Parkhurst, S.M. (2004). *Drosophila* Topors is a RING finger-containing protein that functions as a ubiquitin-protein isopeptide ligase for the hairy basic helix-loop-helix repressor protein. J Biol Chem *279*, 17126-17133.

Sivasankaran, R., Vigano, M.A., Muller, B., Affolter, M., and Basler, K. (2000). Direct transcriptional control of the Dpp target omb by the DNA binding protein Brinker. EMBO J *19*, 6162-6172.

Smith, S.T., Wickramasinghe, P., Olson, A., Loukinov, D., Lin, L., Deng, J., Xiong, Y., Rux, J., Sachidanandam, R., Sun, H.*, et al.* (2009). Genome wide ChIP-chip analyses reveal important roles for CTCF in *Drosophila* genome organization. Dev Biol *328*, 518-528.

Soshnev, A.A., Li, X., Wehling, M.D., and Geyer, P.K. (2008). Context differences reveal insulator and activator functions of a Su(Hw) binding region. PLoS genetics *4*, e1000159.

Spana, C., Harrison, D.A., and Corces, V.G. (1988). The *Drosophila melanogaster* suppressor of Hairy-wing protein binds to specific sequences of the *gypsy* retrotransposon. Genes Dev *2*, 1414-1423.

Spilianakis, C.G., and Flavell, R.A. (2004). Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. Nat Immunol *5*, 1017-1027.

Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes & development *20*, 2349-2354.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Struhl, K. (1984). Genetic properties and chromatin structure of the yeast gal regulatory element: an enhancer-like sequence. Proc Natl Acad Sci U S A *81*, 7865-7869.

Thorvaldsen, J.L., Duran, K.L., and Bartolomei, M.S. (1998). Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. Genes Dev *12*, 3693-3702.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell *10*, 1453-1465.

Udvardy, A., Maine, E., and Schedl, P. (1985). The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. J Mol Biol *185*, 341-358.

Urata, Y., Parmelee, S.J., Agard, D.A., and Sedat, J.W. (1995). A three-dimensional structural dissection of *Drosophila* polytene chromosomes. J Cell Biol *131*, 279-295.

Valenzuela, L., and Kamakaka, R.T. (2006). Chromatin insulators. Annu Rev Genet *40*, 107-138.

Volpi, E.V., Chevret, E., Jones, T., Vatcheva, R., Williamson, J., Beck, S., Campbell, R.D., Goldsworthy, M., Powis, S.H., Ragoussis, J.*, et al.* (2000). Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. J Cell Sci *113 (Pt 9)*, 1565-1576.

Wallace, H.A., Plata, M.P., Kang, H.J., Ross, M., and Labrador, M. (2009). Chromatin insulators specifically associate with different levels of higher-order chromatin organization in *Drosophila*. Chromosoma.

Wallace, J.A., and Felsenfeld, G. (2007). We gather together: insulators and genome organization. Curr Opin Genet Dev *17*, 400-407.

Walters, M.C., Fiering, S., Eidemiller, J., Magis, W., Groudine, M., and Martin, D.I. (1995). Enhancers increase the probability but not the level of gene expression. Proc Natl Acad Sci U S A *92*, 7125-7129.

Wang, Q., Carroll, J.S., and Brown, M. (2005). Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. Mol Cell *19*, 631-642.

Wei, W., and Brennan, M.D. (2001). The *gypsy* insulator can act as a promoter-specific transcriptional stimulator. Molecular and cellular biology *21*, 7714-7720.

Weiler, K.S., and Wakimoto, B.T. (1998). Chromosome rearrangements induce both variegated and reduced, uniform expression of heterochromatic genes in a development-specific manner. Genetics *149*, 1451-1464.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T.*, et al.* (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796-801.

West, A.G., Gaszner, M., and Felsenfeld, G. (2002). Insulators: many functions, many mechanisms. Genes Dev *16*, 271-288.

Yusufzai, T.M., and Felsenfeld, G. (2004). The 5'-HS4 chicken beta-globin insulator is a CTCF-dependent nuclear matrix-associated element. Proc Natl Acad Sci U S A *101*, 8620-8624.

Zhao, K., Hart, C.M., and Laemmli, U.K. (1995). Visualization of chromosomal domains with boundary element-associated factor BEAF-32. Cell *81*, 879-889.

Zhou, J., Barolo, S., Szymanski, P., and Levine, M. (1996). The Fab-7 element of the bithorax complex attenuates enhancer-promoter interactions in the *Drosophila* embryo. Genes Dev *10*, 3195-3201.

VITA

Hyuck-Joon Kang attended Whi-Moon High School, Seoul, Korea. In 1990 he entered Hanyang University in Korea. He served mandatory military service in Korea from 1992 to 1995, and received the degree of Bachelor of Science in 1997. After graduation, he was employed as an academy-industry cooperative research assistant at the Korea Food & Drug Administration (KFDA) and also continued study in graduate school at Korea University. In August 1999, he received the degree of Master of Science. He joined as a research assistant in the Department of Microbiology in Hanyang Medical School. In August 2003, he started his Ph.D training in Dr. Mariano Labrador's laboratory at The University of Tennessee, Knoxville. He was honored with the Cokkinias Graduate Student Excellence Award in 2009.